Applications of Proton Transfer Reaction and Selected Ion Flow Tube Mass Spectrometry in Health Monitoring

Thesis

Applications of proton transfer reaction and selected ion flow tube mass spectrometry in health monitoring

Célia Maria Farinha Lourenço

BSc, MSc

October 2016

A thesis submitted to the Open University in the subject of Analytical Chemistry for the degree of Doctor of Philosophy.

Department of Life, Health and Chemical Sciences

Open University

'Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.'

Marie Skłodowska-Curie (1867 - 1934) chemist & physicist, Nobel Prize for Physics in 1903, in *Our Precarious Habitat* (1973) by Melvin A. Benarde.

Abstract

This thesis investigates the use of Volatile Organic Compounds (VOCs) in disease diagnosis and monitoring. VOCs may be found in the human body, in exhaled breath, faecal matter, urine, and skin. Analysis of the volatile profile produced in the human body can provide an indicator of metabolic status, allowing the screening and monitoring of different diseases and conditions, non-invasively and painlessly.

In this thesis a range of highly sensitive analytical techniques have been adopted to measure such VOCs and demonstrate that such monitoring may be used as a disease diagnostic. For example breath samples may be analysed and calibrated against gas-phase standards prepared under physiologically representative concentrations as a tool for non-invasive disease monitoring, e.g. type 2 diabetes.

Detailed faecal headspace analyses of two different mouse models of type 2 diabetes (Cushing´s mice and *Afmid*) were made. The mouse model of Cushing's syndrome develop excessive circulating glucocorticoid concentrations, which are associated with obesity, hyperglycaemia and insulin resistance. The *Afmid* knockout mice suffer inactivation of *Afmid* genes, which in part regulates many functions including pancreatic secretion. These mice show impaired glucose tolerance. The gut microbiota of diabetic mice appear to have a different composition when compared to wild-type littermates, i.e. significantly increased levels of short-chain fatty acids (SCFAs), ketones, alcohols and aldehydes were found in the faecal headspace of diabetic mice, and a possible link between gut microbiota and type 2 diabetes is demonstrated.

The use of VOCs as a screening tool of colorectal cancer was also explored. The current screening tools show lack of sensitivity and specificity for the screening of the disease. The volatile faecal profile of patients with colorectal cancer was investigated, and

sulphide compounds, including hydrogen sulphide (H_2S) are shown to have potential as biomarkers for screening of colorectal cancer.

Acknowledgments

This work was done due to the collaboration of several people, which have contributed in so many different ways for the scientific achievement that I set out to myself. First of all I would like to acknowledge the confidence that had been put on me to undertake such remarkable research funded by the European Commission, in particular the Marie Curie Fellowship awarded to myself (Marie Curie Actions— Initial Training Networks (ITN), FP7-PEOPLE-2011-ITN, PIMMS (Proton Ionisation Molecular Mass Spectrometry) project 287382).

I would like to thank to my supervisor Prof Claire Turner for all the scientific knowledge she always tried to pass to me; the encouragement to overcome technical hitches and to achieve trustworthy and successful outcomes; the genuine concern of ensuring everything was running well with me, including the life-style in a foreign country. To my supervisor Prof Nigel Mason, thank you for all the feedback given and the supportive thoughts to improve my research.

Thank you to Prof Roger Cox for kindly welcoming me into MRC Harwell in Oxford, and for giving me the opportunity to work in its facilities. To Dr Liz Bentley and Dr Marianne Yon I am sincerely grateful for all the assistance in the experimental setup using laboratory animals. A special thanks to Lucie Vizor, MRC Harwell technician, for taking care so well of my chubby mice.

Within the network, I had the opportunity to share and learn numerous skills. At Kore Technology I met Dr Fraser Reich and the remaining working team, and I am sincerely grateful for all the helpful insights and training on PTR-MS. Prof Chris Mayhew and Ramón, thank you for letting me work on PTR-MS at University of Birmingham. Thank you

to Dr Wolfram Miekisch at University of Rostock and Dr Jan Rozman at Helmholtz Zentrum Munchen for the training given at their institutions.

Julia Barkans and Brett Keith, a special thanks to all the support given at OU labs, as well as Simona Nicoara and Imran, all the training and recommendations that helped me to boost the development of skills in gas chromatography.

To all my dearest friends Claire Batty, Heli Brahmbhatt, Dušan, Raquel, and Prema, thank you for all the brainstorming, all the laughs and good time we had together.

The adaptation to the UK life-style was not always easy. To my long-distance Portuguese friends who were always there to support me and cheer me up, Ana Lucena and Elsa Mora, a big thank you.

Thank you to my family for understanding and supporting this lifetime choice I made in travelling abroad looking for a new professional achievement. *Obrigado à minha família por entender e apoiar esta escolha de vida que eu fiz em viajar para o estrangeiro em busca de uma nova realização profissional.*

To my wonderful boyfriend, fiancé, and long last husband Rui which genuinely supported and encouraged me throughout this journey, who always believed and pushed me to keep going even through the hardest times. *Obrigada meu amor.*

Table of Contents

Abstract……5

Acknowledgments……7

List of Figures……12

List of Tables……16

List of abbreviations and acronyms……18

Chapter 1. The use of volatile organic compounds in disease diagnosis and monitoring

1. 1 Introduction……24
1. 2 The use of VOCs as disease biomarkers……26
1. 3 VOC production in the human body and their use as biomarkers……27
1.3.1 VOC detection in human breath……28
1.3.2 Influence of age/ gender and ovulation on VOC concentrations in human breath……36
1.3.3 Influence of diet on VOC concentrations in human breath……37
1.3.4 Influence of smoking and background contaminants of breath analysis measurements……38
1.3.5 Influence of physiological parameters, physical activity, and stress on VOC concentrations in human breath……39
1. 4 Collection and storage of volatile samples……42
1. 5 Diabetes mellitus……50
1.5.1 Type 1 Diabetes……51
1.5.2 Type 2 Diabetes……51
1.5.3 Regulation of blood glucose levels……52
1.5.4 Methods of glucose determination and monitoring……53
1.5.5 Volatile metabolites as monitor of diabetes……54
1. 6 Mouse models of Type 2 Diabetes……57
1.6.1 Mouse model of Cushing´s Syndrome……58
1.6.2 Afmid mice……60
1.6.3 VOC analysis in rodent models……61
1. 7 Colorectal cancer……63
1.7.1 Screening Programmes and Strategies……64
1.7.2 Use of volatile biomarkers in screening for colorectal cancer……65

Chapter 2. Analytical methods for analysing VOCs

2. 1 Gas-phase ion chemistry……68
2. 2 Techniques for VOC measurements……69
2. 3 Mass Spectrometry and Gas-phase ion chemistry……71
2.3.1 Ionization methods……72
2.3.2 Gas-phase ion-molecule reactions……74

2.3.3 Types of ion-molecule reactions 75
2.3.4 Proton transfer 76
2.3.5 Theoretical prediction of proton transfer rate coefficients 82
2. 4 SIFT-MS 86
2.4.1 Quantitative determination of VOC concentration in SIFT-MS 90
2. 5 PTR-MS 91
2.5.1 Theoretical determination of VOC concentration in PTR-MS 95
2. 6 Gas chromatography-mass spectrometry combined with thermal desorption 97
2.6.1 Calibration with internal standards 99
2.6.2 Desorption of VOCs from TD tubes 100
2.6.3 Chromatographic separation 101
2. 7 Biomarkers vs. Biomarker Profiles 102

Chapter 3. Materials and analytical/statistical methods

3. 1 A potential method for comparing instrumental analysis of breath gas volatile organic compounds using standards calibrated for the gas phase 106
3.1.1 Samples 106
3.1.2 SIFT-MS 110
3.1.3 PTR-MS 111
3.1.4 TD-GC-MS 111
3.1.5 Data processing 113
3.1.6 Product ions and branching ratios 114
3. 2 Analysis of VOC profile in faecal headspace, and exhaled breath and skin from type-2 diabetic mice 117
3.2.1 Mouse model of Cushing´s Syndrome 119
3.2.2 Single Afmid knockout mice 119
3.2.3 Metabolic cages 119
3.2.4 Blood and plasma biochemistry 120
3.2.5 Offline breath- and skin analysis in unrestrained mice 121
3.2.6 Faecal headspace analysis 123
3. 3 Analysis of the volatile faecal metabolome in screening for colorectal cancer 124
3.3.1 Sample acquisition 124
3.3.2 Ethics statement 125
3.3.3 VOC sampling and analysis 125
3. 4 Univariate Analysis 126
3.4.1 Kolmogorov-Smirnov statistic and Shapiro-Wilk statistic test 126
3.4.2 T-test 127
3.4.3 One-way repeated measures ANOVA with Bonferroni post-hoc test 127
3.4.4 Mann-Whitney U test 127
3.4.5 Spearman´s rho correlation coefficient 128
3.4.6 Friedman test (repeated measures) 128
3.4.7 Wilcoxon Signed Rank test using a Bonferroni post-hoc test 129
3. 5 Multivariate statistics 129
3.5.1 Discriminant analysis 129
3. 6 Sensitivity and Specificity 130

Chapter 4. A potential method for comparing instrumental analysis of breath gas volatile organic compounds using standards calibrated in the gas phase

4. 1 Product ions and branching ratios 133
4. 2 Theoretical prediction of reaction rate coefficients 134
4. 3 Quantitative determination of VOC concentrations 138
4.3.1 Individual headspaces 138
4.3.2 TD-GC-MS results 145
4.3.3 VOC mixtures 147

Chapter 5. A longitudinal study of the VOC profile emitted by the faecal headspace of two different mouse models of diabetes

5. 1 A longitudinal study of the VOC profile emitted by the faecal headspace of Cushing´s syndrome mouse model of type 2 diabetes 152
5.1.1 Univariate Analysis using SIFT-MS data 157
5.1.2 Multivariate Statistics using SIFT-MS data 176
5.1.3 Univariate Analysis using GC-MS data 178
5. 2 A longitudinal study of the VOC profile emitted by the faecal headspace of single Afmid knockout mice exhibiting impaired glucose tolerance 185
5.2.1 Univariate Analysis using SIFT-MS data 190
5.2.2 Multivariate Statistics using SIFT-MS data 197
5.2.3 Univariate Analysis using GC-MS data 199

Chapter 6. Analysis of the volatile faecal metabolome in screening for colorectal cancer

6. 1 Introductory notes 206
6. 2 Univariate Analysis 206
6. 3 Multivariate Statistics 211

Chapter 7. Conclusions and future work

7. 1 Conclusions 214
7. 2 Future work 223

References 225

Appendix A. ELISA Assay 239

Appendix B. Lourenço, C.; Turner, C., Metabolites, 2014, 4, 465-498 241

Appendix C. Batty, C. A.; Cauchi, M.; Lourenço, C.; Hunter, J. O.; Turner, C., PLOS ONE, 2015, 10 (6), e0130301 275

List of Figures

Figure 1. 1 – Schematic representations of the steps involved in the identification of VOC as disease biomarker. 26
Figure 1. 2 – **(a)** Schematic drawing of a normal capnogram and typical modes of sampling: **Phase I** is the end of inspiration and beginning of expiration. Gas sampled during this phase represents anatomical dead space air and would typically not contain CO_2 and endogenous VOCs. **Phase II** reflects the appearance of CO_2 and a steep upstroke of CO_2 tension in the normal capnogram. Gas sampled during this phase typically contains a mixture of alveolar and dead space air. **Phase III** reflects the alveolar or expiratory plateau. As the result of alveolar emptying, $PetCO_2$ represents the end-tidal concentration. **(b)** Alveolar breath sampling at the bed-side. Alveolar gas samples are withdrawn from the circuit under visual control of expired CO_2 (Miekisch and Schubert, 2006). 33
Figure 1. 3 – Surface model for common adsorbents. **(A)** Carbotrap surface area ca. 100 m^2/g, uniform charge distribution over all carbon atom centres; **(B)** Tenax (2,6-diphenyl-p-phenylene oxide), surface area ca. 35 m^2/g, non-uniform charge distribution, the charge is essentially localised on the oxygen atoms; **(C)** Amberlite XAD-2, surface area ca. 300 m^2/g, non-uniform charge distribution, less polar than Tenax (Hübschmann, 2009). 46
Figure 1. 4 – An example of a stainless-steel TD tube (Perkin-Elmer) and the active sampling direction (Adapted from Hübschmann, 2009). 48
Figure 1. 5 – Diagram showing insulin action on adipose tissue and the liver to regulate blood glucose levels. 51
Figure 1. 6 – Chemical formula of N-ethyl-N-nitrosourea (ENU). Ethyl group is highlighted in red. 57
Figure 1. 7 – Alkylation of thymine results in the formation of O4-Ethylthymine which is recognised as cytosine and mispairs with guanine (Adapted from Noveroske et al., 2000) 58
Figure 1. 8 – A schematic representation of brain-gut dependence, and brief description of serotonin synthesis from tryptophan (O'Mahony et al., 2015) 59

Figure 2. 1 – Components of a mass spectrometer 70
Figure 2. 2 – Electron ionization (EI) source showing the generation of positively charged ions towards the mass analyzer (Sparkman et al., 2011) 73
Figure 2. 3 – Potential energy diagram for a substitution reaction in the gas-phase and in solution in water (Hoffman and Stroobant, 2007) 74
Figure 2. 4 – A schematic of the selected ion flow tube apparatus, SIFT (Smith and Španěl, 2015). 87
Figure 2. 5 – A schematic of the Series I KORE PTR-MS used for this work equipped with a reflectron TOF-MS. Note that in this schematic the drift tube is not equipped with an ion funnel (Adapted from Kore Technology). 93
Figure 2. 6 – A schematic of the Series I KORE PTR-TOF-MS used for this work. Features of the PTR ion source adapted with a hollow cathode discharge ion source containing two electrodes, one cathode negatively charged and anode positively charged (Adapted from Kore Technology). 93
Figure 2. 7 – A schematic representation of a thermal desorption GC-MS system (Adapted from Poole, 2012). 98
Figure 2. 8 – Calibration Solution Loading Rig (CSLR™), Markes International Ltd (MARKES, 2014). 99
Figure 2. 9 – ULTRA™ thermal desorption autosampler, MARKES International. 100

Figure 3. 1 – Mk2 instrument SIFT-MS (PDZ Europa, UK) and Agilent 6890/5973 GC-MS system equipped with a Markes TD autosampler, and Markes UNITY thermal desorber. 112

Figure 3. 2 – Metabolic cage Techniplast Kettering, UK (Adapted from TECNIPLAST). 120
Figure 3. 3 – Direct desorption into SIFT-MS using a prototype thermal desorber developed by Kore Technology Limited. 122
Figure 3. 4 – Illustrative diagram showing the methodology used for the analysis of VOCs via SIFT-MS and TD-GC-MS.....123

Figure 4. 1 - Calibration curves for 1-propanol, 2-propanol, ethanol and methanol at physiologically representative concentrations; data acquired using SIFT-MS **(a)** and PTR-MS **(b)**. 141
Figure 4. 2 - Calibration plot for acetone **(a)** and acetaldehyde **(b)** at physiologically representative concentrations; data acquired using SIFT-MS and PTR-MS. 142
Figure 4. 3 - **(a)** Concentration of acetaldehyde in ppbv and (**b)** 2-propanol in ppbv, calibrated to 5 ppm in the gas-phase, and normalised counts per second (ncps) as a function of reduced electric field strength (E/N = 90, 125, 160); **(c)** Acetaldehyde-H_3O^+ reaction time and reaction rate coefficient as a function of E/N = 90, 125, 160. 144
Figure 4. 4 - Calibration functions using the internal standard procedure for 1-propanol, 2-propanol **(a)** and acetone **(b)**; data acquired using TD-GC-MS. 145
Figure 4. 5 - Calibration functions using the internal standard procedure for 1-propanol and 2-propanol within a mixture of VOCs **(a)** and acetone **(b)**; data acquired using TD-GC-MS. 147
Figure 4. 6 - Comparison of the peaks areas observed in chromatograms acquired using TD-GC-MS for individual calibration solutions, and peak areas acquired for a VOC mixture containing the six VOCs analysed **(a)** acetone, **(b)** 1-propanol, and **(c)** 2-propanol; **(d)** comparison of peak areas within the VOC mixtures at gas-phase concentrations of 0.1 ppm, 1 ppm, 5ppm, and 10 ppm. ..148

Figure 5. 1 – Body weights along the age of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using T- test ($p < 0.05$). 153
Figure 5. 2 – Body weights along the age of male and female Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using one-way repeated measures ANOVA with Bonferroni post-test. 154
Figure 5. 3 – Blood glucose concentrations along the age in Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using T- test ($p < 0.05$). 155
Figure 5. 4 – Blood glucose concentrations along the age of male and female Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using Mann Whitney's test ($p < 0.05$). 155
Figure 5. 5 – Plasma insulin concentrations of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Plasma insulin concentrations were taken once the animals were culled just after 20 weeks of age. Data analysed using Mann Whitney's test ($p < 0.05$) 156
Figure 5. 6 – Illustrative mass spectrum of a faecal headspace sample from a mutant (het) Cushing´s mouse (animal ID 4.1h). Data acquired by SIFT-MS H_3O^+, m/z range 10-140, acquisition time of 5 seconds and six iterations. 158
Figure 5. 7 – Boxplots of m/z 57 **(a)**, m/z 93 **(b)**; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 160
Figure 5. 8 – Boxplots of m/z 61 **(a),** m/z 79 **(b)**, m/z 97 **(c)** are likely to be acetic acid; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 163
Figure 5. 9 – Boxplots of m/z 47 **(a),** m/z 65 **(b)**, m/z 83 **(c)** are likely to be ethanol; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 164

Figure 5. 10 – Boxplots of m/z 77; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 164
Figure 5. 11 – Boxplots of m/z 89 **(a),** m/z 107 **(b)**; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 165
Figure 5. 12 – Boxplots of m/z 43; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 166
Figure 5. 13 – Boxplots of m/z 71; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 169
Figure 5. 14 – Boxplots of m/z 88; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 170
Figure 5. 15 – Boxplots of m/z 90; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 170
Figure 5. 16 – Boxplots of m/z 104; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 171
Figure 5. 17 – Boxplots of m/z 77; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 172
Figure 5. 18 – Boxplots of m/z 43; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 173
Figure 5. 19 – Boxplots of m/z 60; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$). 174
Figure 5. 20 – Boxplots of m/z 74; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 174
Figure 5. 21 – Boxplots of m/z 78; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 175
Figure 5. 22 – Boxplots of m/z 61; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 175
Figure 5. 23 – Boxplots of m/z 92; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$)..... 176
Figure 5. 24 – Illustrative chromatogram of a mutant (het) animal at 8 weeks of age **(a)**; zoom in overlapping chromatograms at 8 weeks and 20 weeks of age highlighted in blue, the same animal ID showing the increase in VOC across the age **(b)**. 181
Figure 5. 25 – Acetic acid **(a)** and propanoic acid **(b)** faecal headspace concentration of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by GC-MS and analysed using Mann Whitney's test ($p < 0.05$)..... 183
Figure 5. 26 – Faecal headspace concentration of 2,3-butanedione **(a)**, propanoic acid, 2-methyl- **(b)**, 1-butanol **(c)**, pentanoic acid **(d)**, butanoic acid, 3-methyl- **(e)**, and butanoic acid, 2-methyl- **(f)**

of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data analysed using Mann Whitney's test ($p < 0.05$). 185
Figure 5. 27 – Body weights along the age of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test. 187
Figure 5. 28 – Body weights along the age of male and female Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using non-parametric Friedman test ($p < 0.001$) together with Wilcoxon Signed Rank Test using a Bonferroni adjusted p value ($p < 0.01$). 188
Figure 5. 29 – Blood glucose concentrations along the age in Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$). 189
Figure 5. 30 – Blood glucose concentrations along the age of male and female Afmid (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$). 190
Figure 5. 31 – Plasma insulin concentrations of Afmid (hom) compared with WT littermates on a C57BL6/NTac background. 191
Figure 5. 32 – Illustrative mass spectrum of a faecal headspace sample from a mutant (hom) Afmid mouse (animal ID 51.1c). Data acquired by SIFT-MS H_3O^+, m/z range 10-140, acquisition time of 5 seconds and six iterations. 192
Figure 5. 33 – Boxplots of m/z 59; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 194
Figure 5. 34 – Boxplots of m/z 77; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 194
Figure 5. 35 – Boxplots of m/z 61; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$). 195
Figure 5. 36 – Boxplots of m/z 66; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$). 197
Figure 5. 37 – Boxplots of m/z 56; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background at 8 weeks of age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$). 198
Figure 5. 38 – Illustrative chromatogram of a faecal headspace sample of a mutant (hom) animal at 8 weeks of age **(a)**; zoom in overlapping chromatograms at 8 weeks and 16 weeks of age highlighted in blue, same animal ID showing the increase in VOC across the age **(b)**. 202
Figure 5. 39 – Faecal headspace concentration of acetic acid **(a)**, propanoic acid **(b)**, acetoin **(c)** along the age; and acetic acid at 12 weeks **(d)**, propanoic acid at 12 weeks **(e)**, acetoin at 16 weeks **(f)** and 1-butanol at 12 weeks **(g)** of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$). 204

Figure 6. 1 – Boxplots for m/z values showing statistical significance, m/z 35 **(a)** and m/z 90 **(b)** using H_3O^+ precursor ion. Data analysed using Mann-Whitney´s test ($p < 0.05$). 209
Figure 6. 2 – Boxplots for m/z 96 showing statistical significance, using NO^+ precursor ion. Data analysed using Mann-Whitney´s test ($p < 0.05$). 211
Figure 6. 3 – Boxplots for m/z values showing statistical significance, m/z 48 **(a)**, m/z 62 **(b)**, m/z 66 **(c)**, and m/z 94 **(d)** using O_2^+ precursor ion. Data analysed using Mann-Whitney´s test ($p < 0.05$). 212

List of Tables

Table 1. 1 – Types of sorbents and its application 48
Table 1. 2 – WHO diagnostic criteria for diabetes and intermediate hyperglycaemia 50

Table 2. 1 – A comparison of the characteristics of the available research techniques in VOC measurements (Lourenço and Turner, 2014)..... 70
Table 2. 2 – Ion-molecule reactions in the gas phase. 75

Table 3. 1 – Henry's law constants at 298 K (k°H), ΔvapH/ R values in K and the derived Henry's law constants at 293 K for acetone, ethanol, methanol, 1-propanol, 2-propanol, acetaldehyde in aqueous solution. Mean values are given for k°H and ΔvapH/ R..... 107
Table 3. 2 – Ionization conditions in the drift tube of PTR-TOF-MS. 111
Table 3. 3 – Risk class applied to histological results. 124

Table 4. 1 - Comparison of the product ions identified by SIFT-MS and PTR-MS and their associated ion branching ratios (percentage in parentheses), respectively, calibrated for 5 ppm in a highly humid headspace, for a series of saturated alcohols (ROH), acetone (RO) and acetaldehyde (RO) present in human breath. Errors in the branching ratios are estimated to be less than 10%. 133
Table 4. 2 - Proton transfer reaction rate coefficients k_{cap} (T_{rot}, KE_{cm}) (10^{-9} cm^3 s^{-1}) between hydronium ion (H_3O^+) and selected VOCs at 373 K, T_{rot} is taken as the drift tube temperature. For comparison, reaction rate coefficients k_L, k_{ADO}, and k_{cap} (T_{eff}) are also reported and expressed in 10^{-9} cm^3 s^{-1}. Also given are their polarizabilities, α, expressed in units of 10^{-30} m^3 and their permanent dipole moment, μ, in Debye, D..... 135
Table 4. 3 - Determination of VOC concentrations calibrated to 5 ppm in the headspace at 20 °C and a comparison between the mass spectrometric techniques used and GC-MS; SIFT-MS and PTR-MS sensitivity in ncps/ppbv within the range 10^1-10^3 ppbv. 140

Table 5. 1 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of Cushing´s mice faecal headspace and its significance levels acquired using H_3O^+ precursor ion. Data were analysed using Mann Whitney´s test ($p < 0.05$). 159
Table 5. 2 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of Cushing´s mice faecal headspace and its significance levels acquired using NO^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$). 168
Table 5. 3 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of faecal headspace in Cushing´s mice and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$). 172
Table 5. 4 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity given for faecal headspace of Cushing´s mice data set using SIFT-MS. The best classification results are highlighted in bold. 176
Table 5. 5 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS H_3O^+ data set. 177
Table 5. 6 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS NO^+ data set. 178
Table 5. 7 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS O_2^+ data set, and combining SIFT-MS $H_3O^+ + NO^+ + O_2^+$. 178

Table 5. 8 – Quantification of top 10 abundant compounds in faecal headspace of Cushing´s mice (8, 12, 16 and 20 weeks of age). Data acquired by GC-MS using internal standard addition of d8-toluene. Retention time (RT) is given in minutes. Highlighted in bold is given the three most abundant VOCs in faecal headspace of Cushing´s mice. 179
Table 5. 9 – Median concentrations expressed in ng L^{-1} for statistically significant VOCs tested over the age (8, 12, 16 and 20 weeks of age) of Cushing´s mice. Highlighted in bold is given the compounds that are statistically significant for all time points. Data acquired by GC-MS and analysed using Mann Whitney´s test ($p < 0.05$).....181
Table 5. 10 – Median values for statistically significant m/z values tested over the age (8 and 16 weeks of age) of faecal headspace Afmid mice and its significance levels acquired using H_3O^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$). 192
Table 5. 11 – Median values for statistically significant m/z values tested over the age (8, 12 and 16 weeks of age) of faecal headspace of Afmid mice and its significance levels acquired using NO^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$). 195
Table 5. 12 – Median values for statistically significant m/z values tested for 8 weeks of age of faecal headspace of Afmid mice and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$).....196
Table 5. 13 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity given for faecal headspace of Afmid mice data set using SIFT-MS. The best classification results are highlighted in bold.....197
Table 5. 14 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS H_3O^+ data set. 198
Table 5. 15 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS NO^+ data set. 198
Table 5. 16 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS O_2^+ data set.....198
Table 5. 17 – Quantification of top 10 abundant compounds in faecal headspace of Afmid mice (8, 12 and 16 weeks of age). Data acquired by GC-MS using internal standard addition of d8-toluene. Retention time (RT) is given in minutes. Highlighted in bold is given the three most abundant VOCs in faecal headspace of Afmid mice.....199
Table 5. 18 – Median concentrations expressed in ng L-1 for statistically significant VOCs tested over the age (8, 12 and 16 weeks of age) of faecal headspace of Afmid mice. Data acquired by GC-MS and analysed using Mann Whitney´s test ($p < 0.05$). 201

Table 6. 1 – Median values of Low and High Risk groups for m/z 48, m/z 62, m/z 66 and m/z 94 and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$).....209
Table 6. 2 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity.....211
Table 6. 3 – Best set of predictors discriminating between Low Risk group and High Risk group using the data set H_3O^+ 212

List of abbreviations and acronyms

Acronym	Definition
ACTH	Adrenocorticotrophic hormone
ADO	Average Dipole Orientation
AFMID	Arylformamidase
AMDIS	Automated Mass Spectral Deconvolution and Identification System
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
A-T	Adenine-thymine base pair
B	Magnetic sector field
BET	Buffered end-tidal
BMI	Body mass index
BVOCs	Biogenic volatile organic compounds
CAS	Chemical abstract service
CGMS	Continuous Glucose Monitoring System
CI	Chemical ionization
CoHb	Carboxyhaemoglobin
CRC	Colorectal cancer
CSLR	Calibration solution loading ring
CRDS	Cavity ringdown spectroscopy
DNA	Deoxyribonucleic acid
E	Electric sector
E/N	Ratio electric field/gas number density
EI	Electron ionization
ELISA	Enzyme-linked immunosorbent assay
ENU	*N*-ethyl-*N*-nitrosourea
ESI	Electrospray ionization
FAB	Fast atom bombardment
Fast GC	Fast gas chromatography
FDA	Food and Drug Administration

Acronym	Definition
FFAR2	Free fatty acid receptor 2
FFAR3	Free fatty acid receptor 3
FOBT	faecal occult blood test
FS	Full scan
FT-ICR	Fourier transform ion cyclotron resonance
FT-MS	Flow Tube Mass Spectrometry
FT-OT	Fourier transform orbitrap
GC	Gas chromatography
G-C	Guanine-cytosine base pair
GC-FID	Gas chromatography flame ionization detection
GC-MS	Gas chromatography mass spectrometry
GDH	Glucose-1-dehydrogenase
GDM	Gestational Diabetes Mellitus
GOx	Glucose oxidase
β-HBA	β-hydroxybutyrate
het	Heterozygous
HFD	High fat diet
hom	Homozygous
IABR	International Association of Breath Research
ID	Identity
IE	Ionization energy
IFG	Impaired fasting glycaemia
IGT	Impaired glucose tolerance
IMS	Ion Mobility Spectrometry
IT	Ion trap
LAS	Laser absorption spectroscopy
LOD	Limit of detection
m/z	Mass-to-charge ratio
MANOVA	Multivariate analysis of variance

Acronym	Definition
MIM	Multiple ion mode
MRC	Medical Research Council
MS	Mass Spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
NHS	National Health System
NIST	National Institute of Standards and Technology
NTME	Needle trap micro extraction
OU	The Open University
PA	Proton affinity
PCA	Principal component analysis
PEG	Polyethylene glycol
PET	Polyethylene terephthalate
PLOT	Porous Layer Open Tubular
PLSDA	Partial least squares discriminant analysis
PTR	Proton Transfer Reaction
PTR-MS	Proton Transfer Reaction Mass Spectrometry
PTR-TOF	Proton Transfer Reaction Time-of-flight
QIT-MS	Quadrupole Ion Trap Mass Spectrometry
Q	Quadrupole
RM3	Rat and Mouse Diet No. 3
RT	Retention time
SCFAs	Short chain fatty acids
SIFT	Selected Ion Flow Tube
SIFT-MS	Selected Ion Flow Tube Mass Spectrometry
SMBG	Self-monitoring of blood glucose
SPME	Solid-phase micro extraction
SPSS	Statistical Package for the Social Sciences
SRI-MS	Selective Reagent Ionization Mass Spectrometry

Acronym	Definition
SVMs	Support vector machines
T1DM	Type 1 diabetes mellitus
T2D	Type 2 Diabetes
T2DM	Type 2 diabetes mellitus
T-A	Thymine-adenine base pair
TD-GC-MS	Thermal desorption gas chromatography mass spectrometry
TD	Thermal desorption
TNT	2,4,6-trinitrotoluene
TOF	Time-of-flight
TOF-MS	Time-of-flight mass spectrometry
Tk	Thymidine kinase
UK	United Kingdom
USA	United States of America
VOC	Volatile organic compound
WCOT	Wall coated open tube
WHO	World Health Organization
WT	Wild-type

Chapter 1

The use of volatile organic compounds in disease diagnosis and monitoring

1. 1 Introduction

The contemporaneous technological advance in analytical techniques allows the measurement of volatile organic compounds (VOCs) emitted from clinical samples such as exhaled breath, urine, blood, serum, sputum and faeces. However, in spite of its advantages, diagnostics based on VOCs profiling is not yet widely used in clinical practice. Volatile organic compounds (VOCs) are carbon-based molecules that are volatile at ambient temperature (Pysanenko *et al.*, 2009). VOC analysis, particularly breath analysis, appears as a promising non-invasive method when compared to the traditional blood analysis, and may potentially be used for the screening of several conditions, such as cancer.

During the last decades, Selected Ion Flow Tube Mass Spectrometry (SIFT-MS), Proton Transfer Reaction Mass Spectrometry (PTR-MS) and Gas Chromatography-Mass Spectrometry (GC-MS) with thermal desorption or solid-phase micro extraction (SPME) have been widely used for medical research. SIFT-MS and PTR-MS analytical techniques have been developed for potential medical applications, by using breath analysis, urine analysis, faecal analysis, *in vivo* human skin studies, and *in vitro* cell cultures. SIFT-MS and PTR-MS have been developed for real-time, on-line detection and quantification of trace gases in air with a high sensitivity and wide dynamic range.

VOCs as disease indicators is being explored in both human/patient studies and in animal/mouse models. This thesis reports the results of new studies in the use of VOCs as markers of disease. This thesis is composed on 7 chapters. Chapter 1 introduces VOCs in disease diagnosis and monitoring; parameters that affect the VOC levels detected in the human body; the methods available for collection and storage of volatile samples; and some of the applications of VOC analysis in disease diagnosis and monitoring. Chapter 2 presents the current methods for analysing VOCs and its gas-phase ion chemistry.

Chapter 3 gathers the analytical methods used in the sampling and analysis of VOCs. Experimental studies are presented in chapters 4, 5, 6, and explore three different aspects of current research.

- **Fundamental chemistry studies.** The lack of standardization between different analytical techniques is still a major challenge in monitoring VOCs in breath analysis. Therefore, chapter 4 discusses a potential method for comparing instrumental analysis of breath gas volatile organic compounds found in breath using standards calibrated in the gas-phase, where theoretical prediction of reaction rate coefficients were determined and quantitative determination of VOC concentrations via Selected Ion Flow Tube Mass Spectrometry (SIFT-MS), Proton Transfer Reaction Mass Spectrometry (PTR-MS) and Gas Chromatography-Mass Spectrometry (GC-MS) combined with thermal desorption was achieved.

- **In the future VOC analysis may be potentially used to monitor progress of disease in humans with type 2 diabetes.**

  In chapter 5 the use of mouse models to identify potential indicators in VOCs emitted from faeces related to changes in blood glucose in type 2 diabetes (T2D) and impaired glucose tolerance is discussed. Mouse models can provide large datasets for metabolic profiling in a short period of time and under controlled conditions.

- **VOCs emitted from human body are potentially biomarkers of disease.**

  In chapter 6 a pilot study is presented on the use of the volatile faecal metabolome in screening for colorectal cancer. Screening for colorectal cancer is extremely invasive, and not very effective, requiring many ultimately unnecessary invasive and

unpleasant diagnostic colonoscopies thus analysis of VOCs in faecal headspace can be potentially used to screen for colorectal cancer, non-invasively and painlessly, and thus reduce the number of unnecessary colonoscopies.

In the final chapter 7 the results obtained in the present work are summarised and some suggestion for future research directions presented.

1. 2 The use of VOCs as disease biomarkers

The VOCs are generated within the body, travel around the body via the blood and then they can cross the alveolar interface and appear in exhaled breath, being measured at trace concentrations in the parts-per-million by volume (ppmv) and parts-per-billion by volume (ppbv) levels or lower (Smith and Španěl, 2007). A detailed review on breath analysis in disease diagnosis has been written by myself and published in *Metabolites* open access journal (Lourenço and Turner, 2014) and it may be find in Appendix B. Analysis of the relative differences of VOCs can, therefore, provide an indicator of metabolic status, allowing a distinction between healthy and diseased state of the body. Thus, these techniques have the potential to detect diseases in their early stages, non-invasively and painlessly.

Michael Phillips has been a pioneering breath researcher for more than 30 years. Anton Amann organized the International Association of Breath Research (IABR), and created the Journal of Breath Research followed by the annual international meetings on Breath Analysis starting in 2004. They have set the research community the challenge of diagnosis of diseases based upon VOCs not only in exhaled breath, as well as urine, faeces and skin. The scientific community is willing and eager to study all the parameters but this presents several difficulties that need to be investigated before such a technique can be

brought into clinical practice, these includes defining the methodology for sampling, storage and analysis of VOCs. Diagnostics based upon VOCs requires a reliable instrument ready for real-time analysis. Data acquisition of the measurements is managed by the software installed in the instrument, namely, specific software for PTR-MS, SIFT-MS, and GC-MS. Normalization and scaling is required prior to statistical analysis, followed by the use of unsupervised [e.g. Principal Component Analysis (PCA)] or supervised (e.g. classification) multivariate statistical methods with further method cross-validation. The process continues with identification of relevant VOCs and VOCs profiles combining a biochemical explanation of the data. The screening test should be able to distinguish between healthy and diseased patients with high sensitivity and specificity (Figure 1.1).

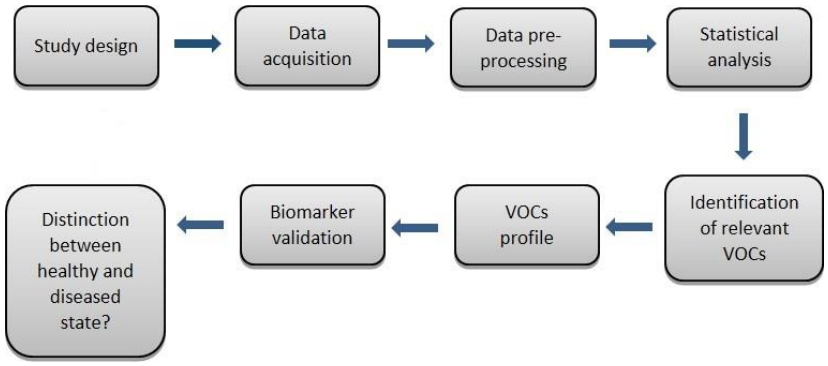


Figure 1. 1 – Schematic representations of the steps involved in the identification of VOC as disease biomarker.

Before exploring further how we may detect VOCs and explore their role as a biomarker for disease it is necessary to understand how and where VOCs are produced in the body.

1. 3 VOC production in the human body and their use as biomarkers

Volatile organic compounds (VOCs) are a diverse group of organic molecules that are volatile at ambient temperature (Pysanenko *et al.*, 2009). At a given temperature, a

substance with higher vapour pressure vaporizes more readily than a substance with a lower vapour pressure. Many of these compounds occur naturally in the environment. Most people encounter VOCs in their day-to-day lives in the form of solvents, constituents of petroleum fuels, etc. In addition, they may be emitted from bodily fluids and as a result, VOCs emitted from urine, faeces, breath, and even skin can provide insight into various biochemical pathways within the body and in the assessment of diseases or drugs monitoring.

The normal range of VOC concentrations largely depends on several factors, such as age; gender; ethnicity; ovulation; environmental contaminants – "*exposome*" – such as radiation, life-style, pollution, diet, and stress; physiological parameters and physical activity. In addition, the type of breath sampling (end-tidal or mixed air) and analysis strongly contribute to the VOC profile obtained.

1.3.1 VOC detection in human breath

VOCs are produced by metabolic processes in various places in the body, e.g. by bacteria in the gut and in the oral cavity by bacterial infections. However, is likely that many of the VOCs which may be detected from the body are not biochemically produced in the body and predominantly come from environmental exposure – so called exogenous VOCs.

VOCs from metabolic processes in the human body may be detected in breath as well as urine and faeces. VOC analysis from the human body started in the 1970s when Linus Pauling and co-workers detected over 200 different VOCs in human exhaled air and in urine headspace by gas chromatography (Pauling *et al.*, 1971). However, many of the VOCs found as major components of breath, such as acetone, isoprene, *etc.* are not of endogenous origin (Miekisch *et al.*, 2004, Smith *et al.*, 2007). A large variety of trace gases

exist in ambient air and these can be taken up via inhalation and skin absorption; the source of others is through ingestion. Hence, some of the trace compounds in exhaled breath, perhaps the majority of them, will be exogenous, and these need to be distinguished from the truly endogenous compounds (Smith *et al.*, 2007, Buszewski *et al.*, 2007).

The endogenous compounds found in human breath, such as inorganic gases (e.g., NO and CO), and VOCs (e.g., isoprene, ethane, pentane, acetone) can be measured directly; other typically non-volatile substances, such as isoprostanes, peroxynitrite, or cytokines, can be measured in breath condensate (Miekisch *et al.*, 2004). These non-volatile substances (isoprostanes, peroxynitrite, or cytokines) are supposed to be present in exhaled breath as aerosol particles.

Background air VOC concentrations are an issue. The concept of alveolar gradient proposed by Michael Phillips (Phillips *et al.*, 1994), defined as *the abundance in breath minus the abundance in room air*, for substances having higher inspired than expired concentrations has been proposed to deal with this, however it, does not properly account for the background air a subject breathes. This has been demonstrated by the work of many researchers, including Schubert (Schubert *et al.*, 2005) and Španěl (Španěl *et al.*, 2013). Indeed, it has been shown (Schubert *et al.*, 2005) that the previous assumption does not lead to quantitative results; instead they suggested that when inhaled (ambient) concentrations of compounds are greater than 5% of the exhaled concentrations, exhaled concentrations cannot be correlated with blood levels with confidence. The levels detected in breath will depend on many factors, including the concentration in ambient air, the duration of exposure, the solubility and partition co-

efficient into tissues, the mass and fat content of the individual, as well as the underlying endogenous concentration. Sample procedures for breath analysis have many advantages over traditional blood analysis of compounds suspended in or dissolved in blood: they are painless and non-invasive, easy to perform, inexpensive, and the results are available immediately for therapeutic assessments (Amann and Smith, 2013). The exhaled air matrix is less complex than that of blood or other body fluids (Buszewski *et al.*, 2007). However, storage of blood is generally easier than breath. Traditional blood analysis typically involves measuring the concentrations of specific salts, proteins or other non-volatile components. However, sampling procedures for blood analysis are stressful for patients and a non-invasive sample, such as breath, is, therefore, preferable.

Mouth- vs. Nose-Exhaled Breath

The challenge of breath analysis for clinical diagnosis and therapeutic monitoring lies in the identification of endogenous volatile compounds present in mouth-exhaled breath which are potential markers of diseases and which reflect levels in the systemic circulation. Many of these trace volatile compounds may be produced in the airways, the oral cavity by bacterial infections, by bacteria in the gut, and also emitted from mucus, saliva and aerosols created in the respiratory tract. Phillips and co-workers performed a pilot study by GC–MS for detecting VOCs in breath associated with oral malodour (Phillips *et al.*, 2005). However, this technique has some limitations since the lower molecular weight VOCs may not be detected due to the sorbent trap selectivity for two carbon atoms or more.

A sampling device of breath exhaled via the mouth, nose, and air in the mouth cavity was developed by Smith and co-workers (Wang *et al.*, 2008). Studies were performed using

SIFT-MS to evaluate the concentration of mouth- and nose-exhaled breath, in order to understand the biological origin of several compounds (Wang *et al.*, 2008, Smith *et al.*, 2008, Pysanenko *et al.*, 2008). Ammonia in the exhaled breath is largely generated in the mouth (Smith *et al.*, 2008) as is ethanol (Wang *et al.*, 2008) and hydrogen cyanide (Wang *et al.*, 2008). This has been demonstrated through showing that the levels measured in the nose exhalations are much lower than those observed in the mouth exhalations. Very low concentrations of propanol and acetaldehyde in exhaled breath appear to be partially systemic and partially mouth generated (Wang *et al.*, 2008). Blood-borne acetaldehyde levels in exhaled breath result from endogenous ethanol metabolism (Lourenço and Turner, 2014). Acetone, methanol, and isoprene showed similar profiles for mouth- or nose-exhaled breath (Wang *et al.*, 2008), indicating that these compounds are totally systemic. However, methanol may be ingested by food intake or drink; hence methanol concentration in breath may not be totally produced through the human biochemistry. The decarboxylation of acetoacetate and the dehydrogenation of isopropanol are the two sources of acetone production in the human body (Kalapos, 2003). The simplicity of the ion chemistry of acetone with H_3O^+ was established in the 90's (Španěl *et al.*, 1997). Thus to avoid the possibility of contamination of the endogenous VOCs through mouth flora, taking breath from the nose is recommended by Smith and co-workers.

Furthermore, a recent study using solid-phase micro-extraction (SPME) of bacterial cultures demonstrated that several compounds detected in mouth-exhaled breath are produced by anaerobic bacteria in tongue biofilms (Khalid *et al.*, 2013b). In addition, poor oral hygiene can be a confounding factor leading to production of ammonia from urea or ethanol from sugars, thus, increasing VOCs concentration in mouth-exhaled breath (Španěl *et al.*, 2006). Ethanol may arise in exhaled breath as an anaerobic fermentation product by gut bacteria. Most breath ethanol, however, appears to be due to mouth

fermentation of sugars (unless the subject has been consuming alcoholic drinks) (Smith *et al.*, 2002).

Additionally, volatile compounds may be produced by bacteria in the gut, transported to and excreted by the lungs (Amann *et al.*, 2010). *Helicobacter pylori* living in the hostile acidic environment of the human stomach release compounds that can be detected in mouth-exhaled air, i.e. converts urea into ammonia, allowing it to survive and grow in gastric acid (Ulanowska *et al.*, 2011). Bacteria present in the gut produce alcohols, including 1-propanol and 2-propanol; structural isomers that exists in the human body (Enderby *et al.*, 2009).

The analysis of mouth- and nose-exhaled breath following ingestion of different doses of alcohol (Smith *et al.*, 2010a) at different concentrations in water was carried out using SIFT-MS by Smith and co-workers. They determined how the volume of ingested liquid influenced the gastric retention and degradation of ethanol. This showed the fraction of ethanol ingested and, consequently, the fraction which enters the blood stream, which in turn is diluted in exhaled breath. The decay of breath ethanol has been followed and measured in mouth- and nose-exhaled breath. Smith and co-workers suggested that saturation of the liver enzymes have an important role in managing the decay of breath ethanol. Additionally, has been shown that gastric retention clearly results in a slower release of ethanol into the gut.

End-tidal and Alveolar breath

The VOCs in breath are at trace levels (ppmv, pptv or lower) and that coupled with the high humidity, means that storage, transport and analysis of samples is challenging due to the possibility of having condensation effects and hydrated hydronium ions acting as precursor ions during analysis. Concentrations of volatile compounds in blood are

reflected by their concentrations in the exhaled air, depending on their blood-gas partition coefficient or solubility. Alveolar breath is the part of exhaled air in equilibrium with systemic blood, whereas end-tidal air is the last fraction of expired air, whose composition resembles alveolar air. Generally, the term end-exhaled breath is applied because it does not imply that the composition of expired air is always identical to the equilibrated air inside the alveoli. It has long been acknowledged that alveolar gas exchange is dependent on ventilation, pulmonary perfusion, and the blood:air partition coefficient (Anderson *et al.*, 2003), thus, there are non-homogeneities in the composition of alveolar air among different lung regions over different blood:air solubilities of VOCs (Levitt *et al.*, 1998). There is evidence that the gas exchange of highly soluble volatile compounds occurs in the airways rather than alveoli (Anderson and Hlastala, 2007), meaning that VOCs measured from the mouth depend on expiratory flow. Generally, the term alveolar breath may be applied for low blood soluble VOCs, whereas for highly soluble volatiles such as acetone, the term end-exhaled breath should be used due to the evidence that gas exchange occurs in the airways rather than alveoli. Such evidence was quantified for the first time by Španěl and co-workers, who demonstrated the discrepancy between the concentration in the alveolar region to that in exhaled air (a factor of 3 for isoprene) (Španěl *et al.*, 2013). Exhaled air is a mixture of dead space and alveolar air (Fowler, 1948). Dead space was previously defined as the volume of expired air, which acts as a conducting airway (nose, pharynx, larynx, trachea), whereas alveolar air is the expired air fraction that has been exchanged in the alveoli. Initial approaches have been taken, such as discarding the first 500 mL of exhaled breath to avoid dilution of the sample by dead space. However, it incorrectly assumes that all subjects have the same volume of dead space (Dubowski, 1974).

Monitoring of expired CO_2 has been used to identify alveolar gas. Thus, Schubert and co-workers measured CO_2 (Schubert *et al.*, 2001, Birken *et al.*, 2006) in the exhaled air of mechanically ventilated patients by means of a capnograph, using a CO_2-triggered alveolar sampling valve (Figure 1.2). They reported this as a reliable and reproducible method for alveolar sampling through CO_2 concentration measures. Later on, Di Francesco and co-workers designed a CO_2-triggered breath sampler suitable for multiple breaths (Di Francesco *et al.*, 2008).

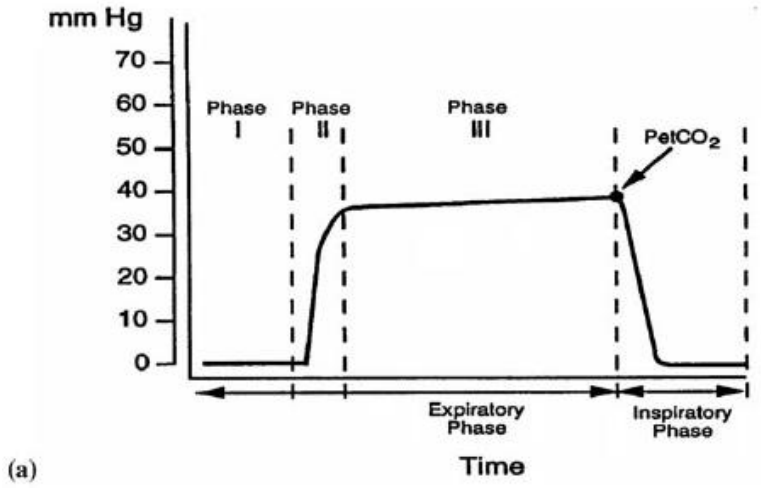


Figure 1. 2 – ***(a)*** *Schematic drawing of a normal capnogram and typical modes of sampling:* ***Phase I*** *is the end of inspiration and beginning of expiration. Gas sampled during this phase represents anatomical dead space air and would typically not contain CO_2 and endogenous VOCs.* ***Phase II*** *reflects the appearance of CO_2 and a steep upstroke of CO_2 tension in the normal capnogram. Gas sampled during this phase typically contains a mixture of alveolar and dead space air.* ***Phase III*** *reflects the alveolar or expiratory plateau. As the result of alveolar emptying, $PetCO_2$ represents the end-tidal concentration.* ***(b)*** *Alveolar breath sampling at the bed-side. Alveolar gas samples are withdrawn from the circuit under visual control of expired CO_2 (Miekisch and Schubert, 2006).*

More recently, Filipiak and co-workers (Filipiak *et al.*, 2012) applied the same on-line monitoring of expired CO_2 in order to collect alveolar air by needle traps used in GC-MS analysis.

The on-line breath sampling so-called buffered end-tidal (BET) (Herbig *et al.*, 2008) breath sampling method has been developed to extend the analysis time of the end-tidal fraction of a single exhalation. This sampling system was designed to buffer only the end-tidal fraction of the breath. The patient is asked to exhale through a tailored tube in which the end-tidal fraction of breath is buffered.

Dilution of VOCs

Hydrophilic exhaled trace gases, such as acetone, interact with the water-like mucus membrane lining the conductive airways, an effect known as wash-in/wash-out behaviour (Anderson *et al.*, 2006). The exhaled breath concentrations of water soluble substances appear to dilute on their way up from the deeper respiratory track to the airway opening, leading to discrepancies between the true alveolar breath and the measured concentrations, demonstrating a dilution effect. It means that highly soluble gases are present in large concentrations in the airway tissue and mucus as compared to less blood-soluble gases for a given partial pressure.

Single or multiple breaths

Breath sampling may be performed for a single breath or for multiple breath cycles (O'Hara *et al.*, 2008). However, the composition of a single breath may not be a representative alveolar gas sample for the reason that breaths may considerably vary from each other due to different modes and depth of breathing. Multiple breaths may be

preferable in order to get reproducible breath samples. However, it may be that exact quantification is less important when characterizing breath profiles.

Comparison of rebreathing and on-line single exhalations of the highly soluble compounds acetone and methanol, and the low soluble isoprene, has been reported previously (O'Hara *et al.*, 2008, King *et al.*, 2012). For highly soluble compounds, such as acetone, exchange occurs in the airways rather than alveoli (Anderson and Hlastala, 2007), thus, breath sampling becomes much more complicated because acetone concentration in end-exhaled breath may not be in equilibrium with the systemic blood. For that reason, an isothermal rebreathing model has been proposed for estimating the alveolar levels of highly soluble exhaled endogenous volatiles (King *et al.*, 2012).

1.3.2 Influence of age/ gender and ovulation on VOC concentrations in human breath

The age and gender of the patient is an important factor in VOCs analysis. PTR-MS was used for determination of isoprene concentrations in the exhaled breath of children by Taucher and co-workers, and was found to be significantly lower than in adults (Taucher *et al.*, 1997). Lechner and co-workers measured the VOCs in the breath of 126 volunteers, using the same technique, reporting an increase in isoprene concentration of exhaled air of male subjects with age (Lechner *et al.*, 2006). SIFT-MS was also used to determine the concentrations of some metabolites in the breath of healthy children aged between 7–18 years old (Enderby *et al.*, 2009); where the median concentration of pentanol (15 ppbv) was also determined. The exhaled breath of several volunteers within the age range 4–83 years was measured and reported a trend of increasing breath ammonia concentration with age (Španěl *et al.*, 2007a). Isoprene is apparently elevated in breath during adolescence, as reported by Smith and co-workers, probably due to the onset of puberty, as stated by the authors (Smith *et al.*, 2010b).

During analysis of the headspace of urine from a number of female volunteers, Smith and co-workers observed acetone and ammonia levels occasionally higher than the normal, i.e. 3- to 12-fold increase in the level of acetone in the urine headspace, and similar increase in the ammonia levels although these were a day out of phase. The urine samples were collected before any food intake. Such findings suggested that it may be caused by metabolic changes occurring during ovulation and related to menstrual cycle length (Diskin *et al.*, 2003, Smith *et al.*, 2006). Studies applied to VOCs present in exhaled breath correlated with ovulation have not yet been reported.

1.3.3 Influence of diet on VOC concentrations in human breath

The levels of breath metabolites are influenced by food intake (Smith and Španěl, 2005, Španěl and Smith, 1999). The breath metabolites ammonia, methanol, ethanol, propanol, formaldehyde, acetaldehyde, isoprene, and acetone were quantified by SIFT-MS for a group of five volunteers, before and after ingesting 75 g of glucose in the fasting state (Turner *et al.*, 2008). Increased levels in the blood/exhaled breath after the consumption of alcohol were also observed, as reported by Smith and co-workers, using SIFT-MS (Smith *et al.*, 2002). Similarly, studies on ethanol metabolism were recently reported by Winkler and co-workers (Winkler *et al.*, 2013). Metabolic degradation of ethanol was tracked by the ingestion of isotope-labelled ethanol using real-time breath gas analysis with PTR-MS. The findings indicated that in part, ethanol was metabolized to acetone and isoprene via acetyl-CoA, as deuterated acetone and isoprene were observed in the mass spectra. Usually, ethanol is metabolized into acetaldehyde in the liver, however, signals from the deuterium-labelled acetaldehyde were not observed, suggesting that this intermediate product was rapidly further metabolized. The VOCs emitted by mouth-exhaled breath after garlic ingestion was assessed by PTR-MS, and the main constituents of garlic were

reported (Taucher *et al.*, 1996). Other products, such as carrots, potatoes, onion, mint, banana and coffee, are also known to emit volatiles at trace concentrations (Sánchez-López *et al.*, 2014, Wang *et al.*, 2016, Alasalvar *et al.*, 2001, Hui, 2010).

1.3.4 Influence of smoking and background contaminants of breath analysis measurements

It has been shown that the composition of exhaled breath is considerably influenced by exposure to pollution and indoor-air contaminants, for instance, smoking-enhanced acetonitrile concentrations were found in the breath and urine of smokers (Jordan *et al.*, 1995, Španěl *et al.*, 1996, Abbott *et al.*, 2003). Compounds present in cigarette smoke, such as 2,5-dimethylfuran, acetonitrile, benzene, toluene, and styrene, can also be identified in smokers and passive smokers' breath (Buszewski *et al.*, 2009, Prazeller *et al.*, 1998). Acetonitrile in breath is a good indicator of whether a given subject is indeed a smoker or not because the concentration of acetonitrile in breath takes nearly a week after cessation of smoking to decrease to that of non-smokers (Jordan *et al.*, 1995). Such findings were reported in an early study using PTR-MS. Smoking increased exhaled ethane and pentane levels in breath. This may be caused by high concentrations of hydrocarbons in cigarette smoke, as well as oxidative damage caused by smoking (Habib *et al.*, 1995). Hydrogen cyanide (HCN), along with acetonitrile and benzene, are known to be present in exhaled breath (Kushch *et al.*, 2008), after analysing the exhaled breath of smokers compared with non-smokers. Measuring carbon monoxide (CO) in exhaled breath is a well-established method used to differentiate between smokers and non-smokers (Sandberg *et al.*, 2011). As a constituent of cigarette smoke, carbon monoxide enters the blood circulation during smoking and forms carboxyhaemoglobin (COHb). The elimination of CO is primarily by respiration, thus, there is a strong correlation between CO in breath and COHb (Sandberg *et al.*, 2011).

Aside from smoking, there are other ways of VOCs entering the body. Breath contains a diverse range of VOCs that can be taken up by the body through inhalation or skin, and, depending on distribution kinetics, may be present in exhaled breath for different periods after exposure. For instance, limonene is found in most air fresheners and cleaning products and is emitted by wooden furniture and floorings and is known to be soluble in blood and adipose tissue and, therefore, has the potential to be taken up by the body during inhalation (Beauchamp, 2011).

Background contaminants (Schubert *et al.*, 2005, Španěl *et al.*, 2013) are an important issue, particularly, when a compound is present in both alveolar breath and inspired air. One approach is to provide a source of purified air (Steerenberg *et al.*, 2000) to inspire during the breath collection and, in this way, it is possible to determine which of the VOCs have originated endogenously. However, different compounds have different wash-out periods from the human breath (Pleil and Lindstrom, 1998), hence, the use of purified air for removal of exogenous compounds does not work consistently for all compounds.

1.3.5 Influence of physiological parameters, physical activity, and stress on VOC concentrations in human breath

Gas exchange (Greger and Windhorst, 1996) during respiration occurs primarily through diffusion. It takes place between the air within the alveoli and the pulmonary capillaries. Nowadays, it is known that exhalation of breath biomarkers may well depend on physiological parameters, such as blood pressure, heartbeat rate and alveolar ventilation (Schubert *et al.*, 2012). Understanding the influence of these factors is an essential requisite for the development of a reliable methodology based on breath volatiles.

Breath gas concentration can then be related to blood concentrations via mathematical modelling. The simplest model relating breath gas concentration to blood concentrations

was developed by Farhi (King *et al.*, 2011). The Farhi equation (equation 1.1), shows that alveolar air concentration, C_A, is proportional to the concentration of VOCs in mixed venous blood (C_V) and depends on the blood:air partition coefficient, $\lambda_{b:air}$, which describes the diffusion equilibrium between capillaries and alveoli, and ventilation-perfusion ratio, V_A/Q_C. Such a ratio ensures that the ideal amount of blood and gas is received by the alveoli for efficient gas exchange. It depends on the alveolar ventilation (V_A) controlling the transport of the VOC through the respiratory track, and cardiac output (Q_c) controlling the rate at which the VOC is delivered to the lungs. The blood:air partition coefficient, $\lambda_{b:air}$, is strongly dependent on temperature ranging from 23 °C in the mouth to 37 °C in the alveoli, affecting soluble gas exchange (Anderson *et al.*, 2003). This coefficient represents the ratio of the concentration in blood to the concentration in the gas phase.

Equation 1.1

$$C_{measured} = C_A = \frac{C_V}{\lambda_{b:air} + \frac{V_A}{Q_C}}$$

For low blood soluble gases ($\lambda_{b:air} \leq 10$) (Anderson *et al.*, 2003) the measured concentration is dependent on the rates at which blood is pumped through the lungs and ventilation, specifically the ventilation-perfusion ratio, V_A/Q_C, where $C_{measured} = C_A$, meaning that low blood soluble VOCs must exchange completely in the alveoli.

Highly soluble VOCs ($\lambda_{b:air} > 10$) (Anderson *et al.*, 2003) tend to be less affected by changes in ventilation and perfusion, however, hydrophilic exhaled trace gases, such as acetone, interact with the water-like mucus membrane lining the conductive airways. The exhaled breath concentrations of these volatiles appear to dilute on their way up from the deeper respiratory track to the airway opening (dilution effect), consequently for

these highly soluble volatiles the concentration measured in exhaled breath is different from the alveolar air concentration, $C_{measured} \neq C_A$. There is also evidence that, with highly soluble volatile compounds, gas exchange occurs in the airways rather than alveoli (Anderson and Hlastala, 2007).

Some studies have been performed assessing the concentration profiles during exercise (King *et al.*, 2010, Schubert *et al.*, 2012, King *et al.*, 2009, Karl *et al.*, 2001, Smith *et al.*, 2013). Recently, the influence of exercise on mouth-exhaled and nose-exhaled breath was further investigated (Smith *et al.*, 2013). Smith and co-workers reported a significant increase of isoprene breath levels after exercise above the respective inactive levels by a factor of 2 or 3 times, which is in agreement with previous findings (King *et al.*, 2010).

An isoprene gas exchange model has been developed and shows a good fit to breath-isoprene levels measured during exercise. The influence of heartbeat rate and breath rate on isoprene breath concentrations have been assessed, where isoprene levels were measured during exercise (Schubert *et al.*, 2012, Karl *et al.*, 2001). Isoprene concentrations showed drastic increase within the initial seconds of exercise, followed by a decline when the heartbeat rate reached the maximum value and respiration rate increased, and lastly at the end of the exercise isoprene concentrations reached similar levels seen at the beginning. This means that the degree of blood-to-air partitioning of isoprene is very sensitive to heart rate. Such measurements demonstrate a relationship between breath rate volume (Vbr), heartbeat volume (HBV), Henry's law constant (H) and temperature (T), seen in Equation 1.2 (Karl *et al.*, 2001).

Equation 1.2

$$C_{A0} = C_{V0} \cdot exp\left(-\frac{1}{HRT} \cdot \frac{Vbr}{HBV}\right)$$

For VOCs such as isoprene, with low solubility in blood and high volatility (when the Henry's law constant is extremely low) a concentration gradient within the lungs is created and governed by the velocity of the bloodstream pumped through the lungs (proportional to heartbeat frequency) and the breathing rate. Namely, with increases in both heart rate and breathing rate, more efficient partitioning of isoprene to breath air is restored. This means that isoprene evaporates efficiently through the transport via the bloodstream to the lungs, hence, $C_{A0} \neq C_{V0}$, and meaning that isoprene venous blood concentration entering the lungs, C_{V0}, is different from isoprene arterial blood concentration leaving the lungs, C_{A0}.

Moreover, measurements taken during sleep showed enhanced blood isoprene concentration due to lower heartbeat rate achieved during the night (Karl *et al.*, 2001).

The influence of stress (Lima *et al.*, 2013) has been acknowledged as an important factor in changing VOC concentrations and may therefore act as confounding factor in the definition of VOCs as disease biomarkers. Additionally, women have demonstrated higher stress-induced levels of cortisol compared to men (Dickerson and Kemeny, 2004).

Stress plays an important role particularly in VOC sampling from mice, since mice become easily stressed under unfamiliar conditions.

1. 4 Collection and storage of volatile samples

The first guidelines concerning sample collection for breath analysis was released in 1999 by the American Thoracic Society for nitric oxide (NO) monitoring in breath (Society, 1999). Updated guidelines were published in 2005 for measurement of NO in mouth- or nose-exhaled breath (Society, 2005). In 2005, recommendations were also published for exhaled breath condensate sampling and analysis (Horváth *et al.*, 2005). Hence,

reproducibility and reliability of sampling methods and analytical measurement procedures continue to be of critical importance.

Collection of VOCs samples may be accomplished through the use of polymer sampling bags, metal canisters, sorbent tubes, and glass vials used in SPME. The stability of selected breath constituents in polymer sampling bags have been previously investigated and assessed by PTR-MS and GC-MS (Mochalski *et al.*, 2013, Beauchamp *et al.*, 2008, Mochalski *et al.*, 2009). Smaller samples are more vulnerable to VOCs losses by permeation. Additionally, the volume of the sample collected affects the stability of the sample, thus, Mochalski and co-workers recommended sample collections as large as possible to prevent background emissions of contaminants (Mochalski *et al.*, 2013).

Lately, needle trap micro-extraction (NTME) (Trefz *et al.*, 2013) combined with GC has been assessed for sample preparation in VOCs analysis. This is a technique similar to SPME but with the advantage of allowing automatic alveolar sampling. The analysis is quite similar to that used in thermal desorption traps. VOCs are thermally desorbed from the needle trap device and separated, identified and quantified by means of two-dimensional gas chromatography combined with MS detector, GC × GC-MS, which has been applied to solve complex problems of separation. Needle traps have offered increased robustness in comparison to SPME, due to the existence of an extraction sorbent packed inside a hypodermic needle rather than supported on a fragile silica fibre that is exposed to the breath matrix during extraction. The influence of humidity, sample volume, and sampling flow has to be thoroughly evaluated in order to be used for pre-concentration of breath volatiles.

Biological intra- and inter-variability among subjects has been introduced as an issue in VOC sampling, particularly exhaled breath. Therefore, breath analysis methods based on monitoring subjects over time may be desirable because they can serve as their own controls (Beauchamp and Pleil, 2013).

Further aspects related to sampling continue to be debated in the scientific community, such as body posture of the subject when providing the breath sample; hyperventilation; breath holding; control of the flow or volume of breath during collection; sampling via nose or mouth; the number of breath samples to be taken to reduce variability (single or multiple breaths); dilution and contamination of the sample; physiological parameters, such as respiratory rate or heart beat rate; alveolar breath or *end-tidal* volume and dead space; the number of subjects per study to avoid "over-modelling"; and direct analysis or sampling for storage (Beauchamp and Pleil, 2013).

Direct/active sampling is preferable to storage for later analysis since, in this way the decomposition of samples or possible loss of compounds by diffusion is avoided. When direct analysis is not possible, the appropriate storage of VOCs is an important issue to consider. This is an important difficulty with breath analysis, there is a frequent need to store samples, because most of the times the patient cannot get near the instrument, and then the subsequent problems of analysis of stored breath samples. In contrast, sampling and storage (prior to analysis) of urine and other body fluids is much easier than breath. Background emission of pollutants, losses by diffusion through the bag or adsorption to the inner bag, and interactions between sample constituents, namely oxidation artefacts of the sample if stored for too long, may irreversibly modify the original sample composition and consequently, mislead the results and assumptions.

In the studies presented in this thesis, direct sampling and further analysis of samples was not possible to perform, therefore VOCs were tested inside Nalophan® bags. **Nalophan® bags** are made from PET-polyethylene terephthalate, are also popular due to its low price, inertness, and relatively high durability. Nalophan® bags are much cheaper than the Tedlar® bags and they can be discarded after a single use, removing the need for bag cleaning or infection control measures.

Previous studies reported the testing of Nalophan® bags for storing tobacco samples, and investigated the factors contributing to decay of gas samples during storage, between 4 and 40 hours after collection (van Harreveld, 2003). Samples remained relatively stable between 4 and 12 hours after sampling. Nalophan® is slightly porous and volatiles are gradually lost over time by diffusion.

Emissions from Nalophan® bags, were assessed 48 hours after filling for storage of volatile sulphur compounds, and none showed emissions of contaminants, thus, all proving to be excellent materials for VOC collection (Mochalski *et al.*, 2009). Nalophan® bags have proven to be much better materials for VOC collection than Tedlar® bags, in terms of contaminants released during storage, sample stability and cost (Ghimenti *et al.*, 2015a).

A study performed by Gilchrist and co-workers investigated the collection and storage of breath samples containing hydrogen cyanide (Gilchrist *et al.*, 2012). Breath was collected into 25 µm thick Nalophan, 70 µm Nalophan and Tedlar bags, at 20 °C or 37 °C. Results showed better correlation between on-line and off-line concentrations for all bag types at 37 °C. Correlation of hydrogen cyanide concentrations in breath samples stored at 37 °C was good up to 24 hours for the 70 µm Nalophan® and Tedlar® bags. Such findings suggested that either would be appropriate to use for collection of breath containing hydrogen cyanide.

Tedlar® (PTFE-polytetrafluoroethylene) **bags** are the most common materials for breath collection (Mochalski *et al.*, 2013) and are available in a range of sizes. Background contaminants released from the bags must be taken into account. Tedlar® bags showed significant emissions of COS and CS_2, especially for black Tedlar bags with COS and CS_2 emissions being seen at up to 7 ppb after three days of storage (Mochalski *et al.*, 2009).
Temperature dependence strongly influences the composition of breath samples in off-line measurements. As the temperature of breath sample falls below body temperature (37 °C), water vapour condenses on the inside of the bag and takes down water soluble compounds. Warming the sample to body temperature will avoid condensation issues and compound losses due to negative temperature gradients (Gilchrist *et al.*, 2012).
The loss of volatile compounds to condensed water in Tedlar bags used for breath sampling has been previously evaluated (Groves and Zellers, 1996), showing differences between dry and wet matrices smaller than 10%. For VOCs with molecular mass above 110 amu, higher losses were detectable (20%–40%) (Mochalski *et al.*, 2013). Thus, as general rule is recommended storing breath samples in pre-conditioned Tedlar® bags up to 6 hours at the maximum possible filling volume.

Metal canisters are essentially stainless-steel vessels where the internal surface has been electro polished, as a result the internal surface of the canisters is chemically inert (Pleil and Lindstrom, 1995). The canisters are extremely robust and samples can be kept in the canisters for a considerable length of time without degradation. However, these canisters are quite expensive and they are not practical when many samples are analysed.

Sorbent tubes are generally made of glass or stainless-steel and contain the desired sorbent suited for the trapping of VOCs at a specific flow rate. It is common practice to

use multi-bed sorbent tubes, up to three sorbents is generally recommended, allowing the broadest volatility range to be analysed with a single tube. The choice of sorbent principally depends upon the volatility (specifically the vapour pressure) of the particular analyte(s). The adsorption material used must have a high thermal stability and must not contribute to the background of the analysis, therefore it should be inert. The expected adsorption and stability properties should not change for a large number of analyses, even on repeated use. The most common materials used in thermal desorption of VOCs, illustrated in Figure 1.3, include porous polymers (e.g. Tenax), graphitised carbon black sorbents (e.g. carbotrap), carbonised molecular sieves (e.g. sulficarb), and zeolite molecular sieves (e.g. carbosieve III) (MARKES, 2013/2014).

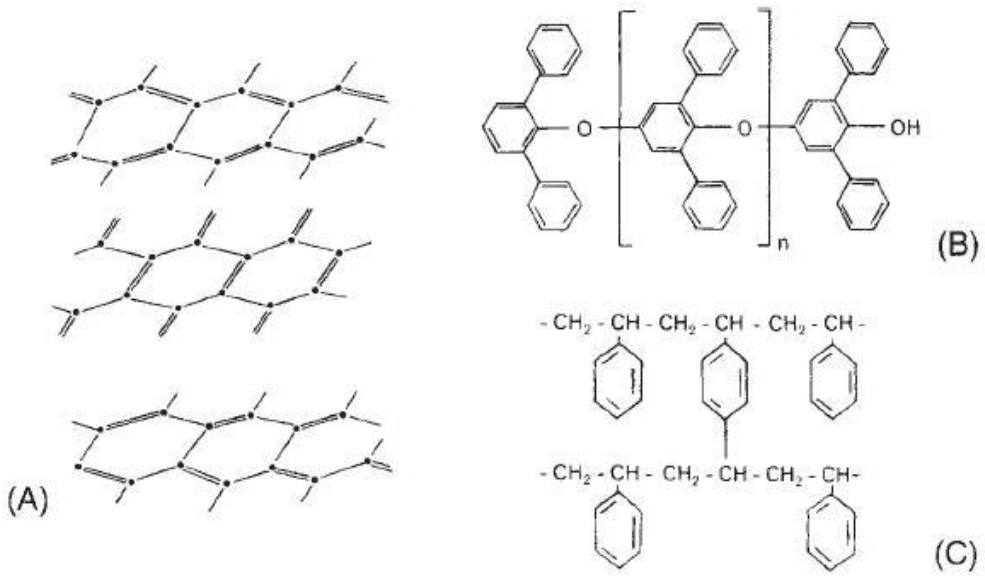


Figure 1. 3 – Surface model for common adsorbents. ***(A)*** *Carbotrap surface area ca. 100 m^2/g, uniform charge distribution over all carbon atom centres;* ***(B)*** *Tenax (2,6-diphenyl-p-phenylene oxide), surface area ca. 35 m^2/g, non-uniform charge distribution, the charge is essentially localised on the oxygen atoms;* ***(C)*** *Amberlite XAD-2, surface area ca. 300 m^2/g, non-uniform charge distribution, less polar than Tenax (Hübschmann, 2009).*

In Table 1.1 the main sorbents currently used for the collection and storage of VOCs are shown. Tenax TA is the worldwide favourite thermal desorption (TD) sorbent due to its specific features, i.e., it is inert and therefore suitable for labile compounds and may be

used for humid atmospheres; the most thermally stable of the porous polymers it also allows efficient desorption at the lowest practical temperature. Carbotrap picks up more volatile compounds and those may be desorbed at higher temperatures. Carbonised molecular sieves such as sulficarb, are the strongest sorbents and are used to trap the most volatile compounds. These sorbents are not very hydrophobic and therefore, may require dry purging prior to desorption to remove excess water. Zeolite molecular sieves are very selective hydrophilic sorbents, thus they provide significant water retention, and are used for specific applications, e.g. occupational hygiene monitoring of nitrous oxide (MARKES, 2013/2014).

Table 1. 1 – Types of sorbents and its application

Sorbent	***Compound range***	***Temperature limits***	***Main features***
Tenax TA	$C_{6/7}$ to C_{30}	280– 320 °C for analysis; 330°C for conditioning	Weak strength; inert and hydrophobic; very efficient desorption of VOCs.
Carbotrap	$C_{3/4}$ to $C_{6/7}$	Up to 360 °C for analysis; 380 °C for conditioning	Weak/medium strength; some activity (e.g. with monoterpenes); hydrophobic; low background.
SulfiCarb	C_3 to C_6	Up to 360 °C for analysis; up to 380 °C during conditioning	Very strong strength; moderate water retention; inert; very volatile compounds and sterically large molecules, e.g. SF_6
Carbosieve III	C_2 to C_5	Up to 360 °C for analysis; up to 380 °C during conditioning	Very strong strength; significant water retention.

Stainless-steel TD tubes are suited for passive or active pumping of a VOC mixture and this is typically performed at flow rates of around 20–100 mL min^{-1} (MARKES, 2013/2014). The VOC mixture is pumped through multiple sorbent beds allowing the trapping of

compounds of different volatilities in a single tube. The initial bed comprises the sorbent of weak strength and at the end the stronger sorbent (Figure 1.4).

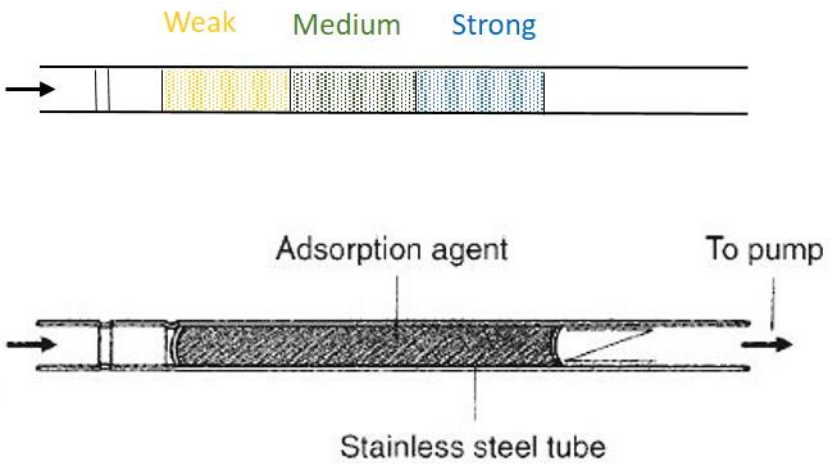


Figure 1. 4 – An example of a stainless-steel TD tube (Perkin-Elmer) and the active sampling direction (Adapted from Hübschmann, 2009).

TD sorbent tubes have been widely used in medical research for the trapping of the headspace of cell cultures (Filipiak *et al.*, 2010); collection of faecal (Proudman *et al.*, 2015) and urine (Thalhamer *et al.*, 2011) headspace; and collection of exhaled breath for exposure (Ghimenti *et al.*, 2015b) and disease (Khalid *et al.*, 2013a) assessment.

Breath analysis was the original way of analysing VOCs from the body and how they differ as a result of disease. However, other samples have also been used in VOC analysis, for example urine and faecal headspace, and VOCs emanating from skin (Batty *et al.*, 2015, Huang *et al.*, 2013, Mochalski *et al.*, 2014a). Some of the applications of VOC analysis in disease monitoring and diagnosis now follow.

1. 5 Diabetes mellitus

Diabetes is a group of metabolic diseases characterized by hyperglycaemia resulting from deficiencies in insulin secretion, insulin action, or both. The chronic hyperglycaemia of diabetes is associated with long-term damage, dysfunction, and failure of different organs, especially the eyes, kidneys, nerves, heart, and blood vessels. According to the World Health Organization (WHO) guidelines there are three variations of Diabetes mellitus: Type 1 diabetes mellitus (T1DM) (insulin-dependent); Type 2 diabetes mellitus (T2DM) (non-insulin dependent); and gestational diabetes mellitus (GDM). The WHO criteria for diabetes and intermediate hyperglycaemia are given in Table 1.2.

Table 1. 2 – WHO diagnostic criteria for diabetes and intermediate hyperglycaemia

	Normal		***Impaired fasting glycaemia (IFG)***		***Impaired glucose tolerance (IGT)***		***Diabetes mellitus (DB)***	
Venous plasma	**Fasting**	**2 h**	**Fasting**	**2 h**	**Fasting**	**2 h**	**Fasting**	**2 h**
Glucose (mmol L^{-1})	<6.1	<7.8	≥6.1 &<7.0	<7.8	<7.0	≥7.8	≥7.0	≥11.1
Glucose (mg dL^{-1})	<110	<140	≥110 &<126	<140	<126	≥140	≥126	≥200

During fasting, the normal levels for a healthy individual are lower than 110 mg dL^{-1}. Impaired fasting glycaemia (IFG), more commonly known as pre-diabetes refers to a condition in which the fasting blood glucose level is consistently elevated above what is considered normal levels. For this condition the levels range between ≥ 110 & < 126 mg dL^{-1} during fasting. IFG can progress to type 2 diabetes mellitus if life-style changes are not made. Impaired glucose tolerance (IGT) is a pre-diabetic state of hyperglycaemia that is associated with insulin resistance and increased risk of cardiovascular pathology. IGT might precede type 2 diabetes mellitus by many years. For impaired glucose tolerance the levels are inferior to 126 mg dL^{-1} during fasting. A person is diagnosed with diabetes mellitus for venous plasma levels equal or higher than 126 mg dL^{-1} for the fasting state.

1.5.1 Type 1 Diabetes

Type 1 diabetes (insulin-dependent) develops when the insulin-producing cells in the body have been destroyed and the body is unable to produce any insulin. The pathogenesis of T1DM remains poorly understood, but the most likely scenario is that an environmental factor triggers a selective autoimmune destruction of the β-cells in the pancreas of a genetically predisposed individual, decreasing the amount of insulin that is produced and resulting in hyperglycaemia. The destruction is irreversible and can take place over a period of several years; this is known as the pre-diabetic phase (Holt and Hanley, 2012). Treatment of T1DM involves exogenous injections of insulin which cannot be taken orally due to insulin degradation in the gastrointestinal tract (Holt and Hanley, 2012).

1.5.2 Type 2 Diabetes

Type 2 diabetes is a heterogeneous disorder creating insulin resistance and insulin deficiency. Type 2 diabetes develops when the body can still make some insulin, but not enough, or when insulin that is produced does not work properly, known as insulin resistance. The World Health Organization states that T2DM comprises 90% of people with diabetes around the world.

The two main pathological components of T2DM are insulin resistance, and β-cell dysfunction and failure. The risk of developing T2DM increases with age, obesity, cardiovascular disease (hypertension) and lack of physical activity (Srinivasan and Ramarao, 2007). Type 2 diabetes is a combination of environmental factors such as diet and lack of exercise, as well as genetic predisposition. It has an onset after the age of 40 years, however due to the increasing prevalence of obesity, it is also being observed in younger people (Song and Hardisty, 2008).

1.5.3 Regulation of blood glucose levels

Insulin is secreted by β-cells of the pancreas, and is crucial to regulating carbohydrate and fat metabolism in the body. Under normal physiological condition, insulin is the principal controller of glucose metabolism, and is triggered by high blood glucose levels. On the other hand, glucagon hormone produced in the pancreas, is triggered by low glucose levels (Holt and Hanley, 2012).

Insulin resistance creates increasingly high levels of hepatic glucose output due to the inability of insulin to produce its usual biological effects, and therefore the build-up of glucose in the blood stream. As a consequence, glucose uptake in adipose tissue decreases. A summary diagram is presented in Figure 1.5. In a vasculature point of view, the consequences of insulin resistance are increased stiffness of arteries and increased coagulability (Holt and Hanley, 2012).

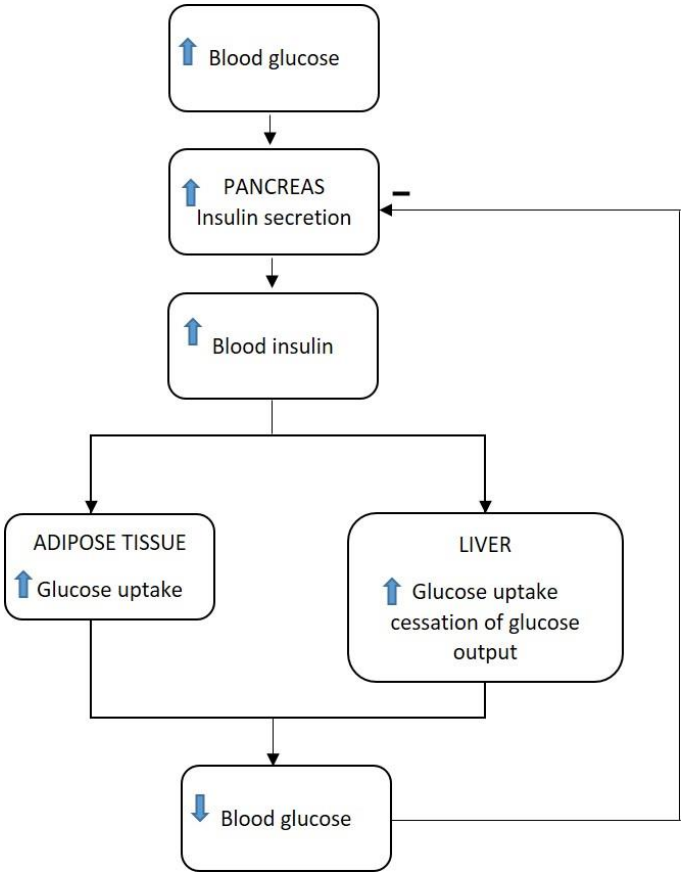


Figure 1. 5 – Diagram showing insulin action on adipose tissue and the liver to regulate blood glucose levels.

The mechanisms leading to the development of insulin resistance are not fully understood. However, there are potential mechanisms of insulin resistance, namely absent or a reduced number of insulin receptors; abnormal insulin receptors; deficiencies in post-receptor signalling.

1.5.4 Methods of glucose determination and monitoring

A key to controlling diabetes is monitoring and controlling blood glucose concentrations. Glucose concentration may be determined in whole blood, plasma, or serum samples. Concentrations in blood are lower than those measured in plasma or serum. This is due to the greater water content of the cellular fraction (Walker *et al.*, 1990). Under usual circumstances, the concentration of glucose in whole blood is about 15% lower than in plasma or serum.

Self-monitoring of blood glucose (SMBG) has been established as a valuable tool for the management of diabetes, assisting in the detection of any incidence of hyperglycaemia or hypoglycaemia, providing real-time information for adjusting medications and dietary regimens. Electrochemical glucose biosensors are mainly preferred for SMBG among the others, optical, thermometric, piezoelectric, and magnetic, due to their better sensitivity, reproducibility, and easy maintenance as well as their low cost (Yoo and Lee, 2010). Enzymatic amperometric glucose biosensors are the most common devices commercially available for self-monitoring of blood glucose. Amperometric sensors monitor currents generated when electrons are exchanged either directly or indirectly between a biological system and an electrode (Heller, 1996, Turner *et al.*, 1999). Glucose biosensors for SMBG are usually based on the two enzyme families, glucose oxidase (GOx) and glucose-1-dehydrogenase (GDH) (Yoo and Lee, 2010). Glucose oxidase is the standard for

SMBG due to its higher selectivity for glucose. Glucose biosensors are reliable, rapid, and accurate.

In vivo glucose monitoring was proposed in 1982 (Shichiri *et al.*, 1982). The first continuous glucose-monitoring device to be approved by the US Food and Drug Administration (FDA) was the Continuous Glucose Monitoring System (CGMS™) (Medtronic MiniMed, Northridge, CA, USA), which consists of a glucose sensor implanted into the subcutaneous tissues and connected to a pager-like monitor worn by the user. The device detects glucose levels in blood every five minutes or automatically overnight (Poscia *et al.*, 2003). However, these devices require calibration against a regular blood glucose meter.

Many diabetics do not test their blood glucose levels periodically because it is unpleasant and inconvenient. Hence it's worth exploring alternative, non-invasive methods of monitoring blood glucose. Other non-invasive alternatives have been studied, namely, saliva (Agrawal *et al.*, 2013), tears and urine (Srinivasan *et al.*, 2003). Just recently, VOCs in breath or skin has been proposed as a potential way to provide information about the glycaemic levels in blood (Righettoni *et al.*, 2013, Turner *et al.*, 2009, Turner *et al.*, 2008).

1.5.5 Volatile metabolites as monitor of diabetes

Centuries ago, John Gallo reported a compound in human breath that had the smell of decaying apples (Crofford *et al.*, 1977). Nowadays, it is well known that the compound was principally acetone (Reichard *et al.*, 1986). Acetone is produced by decarboxylation of acetoacetate and through dehydrogenation of isopropanol (Kalapos, 2003). Diabetic patients exhibit increased concentrations in blood and urine of ketone bodies, acetone,

acetoacetic acid, and beta-hydroxybutyric acid. Ketone bodies are produced by the liver during fatty acid metabolism, and are used as an energy source, instead of glucose, when glucose is not readily available (Smith *et al.*, 2011).

Acetone is a highly soluble compound present as a major common breath metabolite in everyone. This VOC has been identified and quantified previously (Smith *et al.*, 2007, Turner *et al.*, 2006a). Elevated breath acetone levels associated with diabetes mellitus have been reported for more than a decade (Kalapos, 2003, Manolis, 1983, Deng *et al.*, 2004). A few studies have been carried out, involving cohorts of patients with diabetes and healthy controls, in whom parallel with blood glucose analysis. Biochemical changes that occur in the diabetic state may be reflected in changes in the profile of VOCs in exhaled breath, however, the high inter-individual variability in breath acetone concentration is not well understood with effects arising from diurnal variability, fasting status, diet, age, and gender.

Acetone levels were previously quantified for healthy volunteers, before and after ingesting 75 g of glucose in the fasting state (Turner *et al.*, 2008). The intake diet seems to significantly influence the levels in breath acetone. Measurements were taken following a ketogenic diet (Musa-Veloso *et al.*, 2002, Španěl *et al.*, 2011) in the exhaled breath of healthy individuals, and for a small group of individuals suffering from diabetes. Results have shown that breath acetone concentrations increased after ingestion of a ketogenic meal, or either following a low carbohydrate diet (Španěl *et al.*, 2011). Smith and co-workers reported that breath acetone increases substantially during fasting when a change takes place from carbohydrate to fat metabolism (Smith *et al.*, 2011). PTR-MS has been used to report a variation on breath acetone with age and fasting state, but no

statistically significant differences between gender or body-mass index (BMI) were found (Schwarz *et al.*, 2009).

Glycaemic control is essential for management of diabetes. At the moment, blood analysis is the faster way to provide information/monitoring of glucose levels in patients with diabetes. However, studies amongst adults and children showed discomfort associated to blood sampling and needle phobia (Simmons *et al.*, 2007, Hamilton, 1995). Only recently, breath analysis appears as a potential non-invasive method for monitoring glucose concentrations in blood (Turner *et al.*, 2009, Guo *et al.*, 2010, Minh *et al.*, 2011). Turner and co-workers have monitored the breath of eight patients with T1DM using a glucose clamp technique. In all patients, the breath acetone declined linearly with blood glucose concentration. Hence, this study indicates that breath acetone does vary as glycaemia and/or metabolic status changes in T1DM. In contrast little is currently known for T2DM, some measurements of breath acetone were taken, nevertheless, with no significant results (Storer *et al.*, 2011). Monitoring and reliable quantification of VOC concentrations within the body of a diabetic patient is not that easy as VOC fluctuate over the day.

Although the findings have pointed to acetone as a potential biomarker of diabetes, there is no simple association of breath acetone concentration and blood glucose concentration. The issue lies in the fact that acetone generation is linked to lipolysis and blood glucose changes (Kalapos, 2003).

Isoprene has also been proposed as a potential indicator of diabetes (Salerno-Kennedy and Cashman, 2005). However, several studies reported no apparent correlation between blood glucose and breath isoprene (Turner *et al.*, 2006b, O'Hara *et al.*, 2009). Methyl nitrate is also suggested to be correlated with blood glucose in insulin dependent

diabetics, though the levels are reported to be lower than acetone or isoprene. Therefore, this might not be a useful compound for monitoring purposes (Smith *et al.*, 2011).

Gas sensors (Righettoni *et al.*, 2013) have been used to assess the relationship between blood glucose of healthy volunteers and breath components. It should also be noted that dogs (Wells *et al.*, 2008) have been shown to be able to detect hypoglycaemic episodes in their diabetic owners through detecting breath or skin odour.

1. 6 Mouse models of Type 2 Diabetes

Animal models allow identification of genes and how alterations to these genes can impact on the pathways of the disease. Thus, mice are used to investigate the function of genes and develop reproducible models in which the phenotype of the animal can be tested. Mouse mutants are important in understanding biological processes and biochemical or cellular pathways in diseases such as T2DM. The physiology shared between mice and humans makes the mouse ideal for modelling complex human diseases and for drug efficacy testing. This is not just because they are extremely similar to humans at the genomic level, but also because the pathophysiology of disease in mice is similar to that of humans.

The Home Office annual statistics 2012 in Great Britain states that the “mouse” is the available model most used in scientific research because mice are quite small which allows low maintenance and costs; they have a fairly short gestation; and they have a short life span allowing investigation of aging phenotypes. Consequently, mouse models can provide large datasets for metabolic profiling in a short period of time and at controlled conditions. Thus, animal models are an important tool to decipher the various sources of VOCs in a controlled environment.

Similar to humans, the phenotype of a mouse model also depends on the genetic background, gender, and age; this means that some mice are more susceptible to obesity and diabetes than others, and insulin resistance manifested at different time points (Ueda *et al.*, 2000, Cheverud *et al.*, 2004).

In order to understand the complex genetics of insulin resistance and β-cell dysfunction, investigators have generated transgenic and knockout mice bearing mutations in genes required for insulin action/or insulin secretion.

Animal models of T2DM include insulin resistance, β-cell dysfunction and impaired glucose homeostasis, and insulin secretion. Animal models develop diabetes either spontaneously or by using chemical, surgical, genetic or other techniques (McMurray and Cox, 2011).

1.6.1 Mouse model of Cushing´s Syndrome

The mouse model of Cushing´s syndrome is a chemically induced T2DM model. The animals are chemically induced by *N*-ethyl-*N*-nitrosourea (ENU) (Hitotsumachi *et al.*, 1985), a highly efficient mutagen used to induce random point mutations triggering a diabetic phenotype. *N*-ethyl-*N*-nitrosourea (ENU) (Figure 1.6) is a powerful and efficient mutagen, i.e. toxic and carcinogenic, and has been found to induce random point mutations in mouse spermatogonial stem cells (Hitotsumachi *et al.*, 1985), meaning that observed phenotypes will be a consequence of single gene effects. The mutagenic effects of ENU depend on the dosage and strain of the animal used (Noveroske *et al.*, 2000).

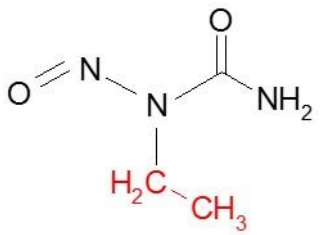


Figure 1. 6 – Chemical formula of N-ethyl-N-nitrosourea (ENU). Ethyl group is highlighted in red.

The mutation is generated by nucleotide substitution. ENU is an alkylating agent and acts by transferring its ethyl group to any of a number of identified nucleophilic nitrogen or oxygen sites on each of the four nucleotides (adenine, thymine, guanine and cytosine) (Noveroske *et al.*, 2000). However, 82% of nucleotide substitutions occur at A-T (adenine-thymine) base pairs to T-A (thymine-adenine), and A-T (adenine-thymine) base pairs to G-C (guanine-cytosine) (Takahasi *et al.*, 2007, Noveroske *et al.*, 2000) (Figure 1.7).

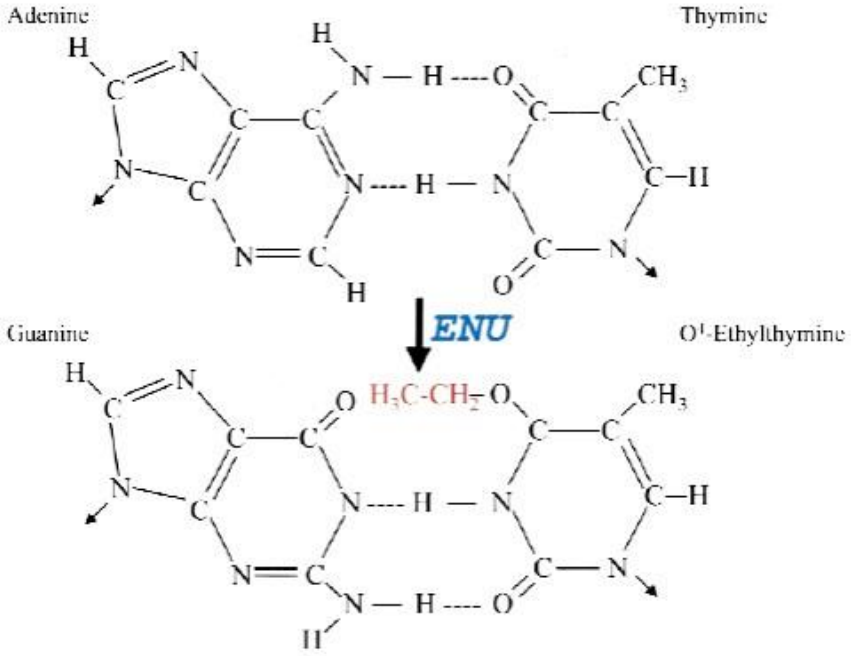


Figure 1. 7 – Alkylation of thymine results in the formation of O4-Ethylthymine which is recognised as cytosine and mispairs with guanine (Adapted from Noveroske et al., 2000)

ENU treatment: A detailed protocol for mouse mutagenesis using ENU is described elsewhere (Salinger and Justice, 2008). ENU is usually administered by a sequence of intraperitoneal injections to adult male mice, followed by a period of infertility due to the depletion of spermatogonia caused by the ENU treatment.

Cushing´s mice develop excessive circulating glucocorticoid concentrations in the body potentially due to adrenocorticotrophic hormone (ACTH) alterations by the anterior pituitary gland, and directly affecting the production and release of cortisol hormone by the cortex of the adrenal gland (Bentley *et al.*, 2014). When a pituitary tumour is the cause of elevated ACTH this is known as Cushing's disease and the symptoms of the excess cortisol is known as Cushing's syndrome. Glucocorticoid is involved in the metabolism of carbohydrates, proteins and fats (Kufe *et al.*, 2003). As a consequence, mice are prone to develop obesity, hyperglycaemia, and insulin resistance, along with hair loss, thin skin and low bone mineral density (Bentley *et al.*, 2014).

1.6.2 Afmid mice

The enzyme arylformamidase (AFMID), also known as kynurenine formamidase participates in tryptophan metabolism (O'Mahony *et al.*, 2015) which in turn regulates many functions including pancreatic secretion, among others. Figure 1.8 illustrates the brain-gut dependence, and brief description of serotonin synthesis from tryptophan.

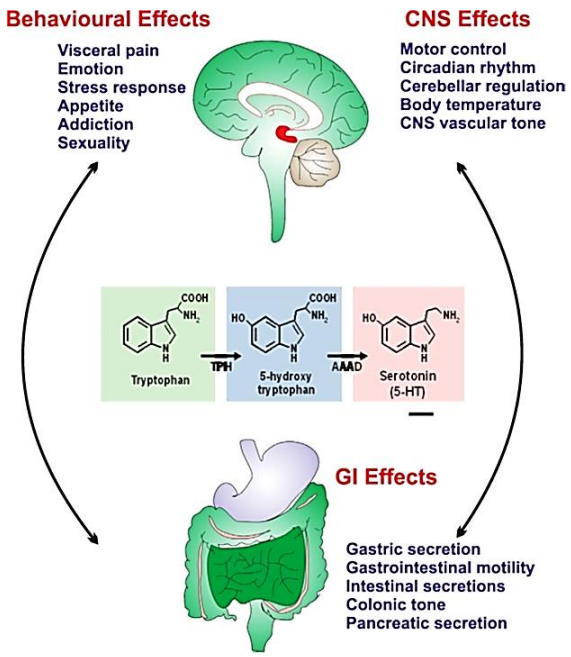


Figure 1. 8 – A schematic representation of brain-gut dependence, and brief description of serotonin synthesis from tryptophan (O'Mahony et al., 2015)

The *Afmid* knockout mice suffer inactivation of *Afmid* genes in order to further investigate possible links between abnormal tryptophan metabolism and diabetes and to examine the effect of single *Afmid* knockout. These mice show impaired glucose tolerance, although their insulin sensitivity is unchanged when compared to wild-type animals. This phenotype results from a defect in glucose stimulated insulin secretion and these mice show reduced islet mass with age (Hugill *et al.*, 2015). Kidney failure has been reported in double (*Afmid* and thymidine kinase (Tk) genes) knockout mice (Dobrovolsky *et al.*, 2005), however no evidence was found for a single *Afmid* knockout, suggesting that this specific phenotype of kidney failure resulted from loss of Tk expression in the double knockout.

The gene knockout was achieved by using Cre-Lox recombination giving deletion of tm1b allele. Cre-Lox recombination is a site-specific recombinase technology, used to carry out deletions, insertions, translocations and inversions at specific sites in the DNA of cells. The system consists of a single enzyme, Cre recombinase, that recombines a pair of short target sequences called the Lox sequences (Liu *et al.*, 2003).

1.6.3 VOC analysis in rodent models

The biochemical origin of VOC in breath, faeces, urine or even skin in rodent models is far from being well understood, furthermore, published literature is scarce. Breath analysis in rodent models has been the preferred method used by researchers to assess the VOC composition within the body. Breath analysis in mice is complex, the sampling is challenging, and the detection of VOC at trace concentration with enough sensitivity is difficult, not to mention the confounding factors needed to be taken into account, i.e.

VOCs arising from other sources, faeces, urine, stress (Lima *et al.*, 2013), and physical activity/movement (King *et al.*, 2009).

Breath acetone has been measured in a rat model to assess the feasibility of using it as an indicator of ketosis induced by high fat, low-carbohydrate ketogenic diet (Likhodii *et al.*, 2002). This ketogenic diet is used primarily to treat difficult-to-control (drug-resistant) epilepsy in children (Lefevre and Aronson, 2000), and it requires testing for its efficacy fairly regularly. Breath acetone was measured and analysed over a thirty three days period and compared against blood glucose and plasma β-hydroxybutyrate (β-HBA), the most commonly measured ketone body. They found a good correlation between breath acetone and plasma β-HBA.

The analysis of exhaled breath in an induced mouse asthma model via ion mobility spectrometry (IMS), anaesthetised prior to breath analysis present many challenges (Vautz *et al.*, 2010). Effects to take into account include contaminants from the background air; contaminants coming from the plastic sampling system; ovulation effects on the measured VOC profile; the influence of anaesthesia on the measured VOC profile.

The online breath gas analysis in real-time in unrestrained mice was established using a setup consisting of a respiratory chamber directly connected to PTR-MS (Szymczak *et al.*, 2014). Later on, a study on the changes of breath VOC profile in diet-induced obese mice was reported (Kistler *et al.*, 2014). More recently, the same research group analysed the breath of unrestrained mice in real-time of two different obese mouse models (Kistler *et al.*, 2016). High fat diet (HFD) induced obese and mono-genetic obese mice, and have looked for alterations in the volatile profile which might be due to either metabolic deregulations, according to the degree of obesity, but could also be due to changes in diet composition.

Other studies have been carried out in other animals, specifically in a caprine animal model (Fischer *et al.*, 2015), where the physiological variability in VOCs in exhaled breath and released from faeces have been evaluated, and the study assessed the effects of feed composition and somatic growth on VOC patterns.

1. 7 Colorectal cancer

Colorectal cancer (CRC) has been attributed to individual genetic predisposition and environmental factors, including life-style and diet. Within life-style factors, elevated body mass index (BMI), obesity, and low physical activity are related to increased risk of colorectal cancer. It has been shown that diet can significantly influence and promote the growth of malignant colon cells, particularly, red meat intake, where protein is the major constituent leading to protein fermentation metabolites (e.g. sulphur compounds, N-nitroso compounds, and phenolic and indolic compounds) potentially carcinogenic and possible linked to colon cancer (Cummings and Bingham, 1987, Van Munster and Nagengast, 1993, Windey *et al.*, 2012, Larsson and Wolk, 2006, Nyangale *et al.*, 2012). Sulphate reducing bacteria in the gut produce sulphur compounds, including H_2S, potentially linked to colon cancer. N-nitroso compounds are known to be carcinogenic and formed via nitrosation of amines. Phenolic (phenol and p-Cresol) and indolic (indole and skatole) compounds are products of catabolism of aromatic amino acids and these have been shown to be toxic to cells (Nyangale *et al.*, 2012).

There is also evidence of a relationship between infective agents and some human cancers. Certain intestinal microbiota play an important role in the pathogenesis of colorectal cancer (Marchesi *et al.*, 2011). Carbohydrate fermentation by bacteria in the gut results in production of hydrogen, methane, carbon dioxide, and short-chain fatty acids (SCFAs) mainly acetate, propionate and butyrate (Lourenço and Turner, 2014).

Significant changes in faecal microbiota, elevated faecal pH, and significant decreases in faecal SCFAs concentrations, have been described in patients with colorectal cancer (Ohigashi *et al.*, 2013). In addition, bacteria have an important role on the generation of the majority of some compounds present in breath, such as hydrogen, hydrogen cyanide, aldehydes, and alkanes (Ajibola *et al.*, 2013).

1.7.1 Screening Programmes and Strategies

Given that colorectal cancer is the second leading cause of cancer-related death in Europe, one of the essential necessities for CRC screening is detection of the disease at earlier stages with high sensitivity and specificity.

In the UK and USA the screening programmes are quite similar and the initial screening stage includes stool/faecal based tests, and then secondary structural based tests. The current chosen faecal based test in the UK's NHS s is the faecal occult blood test (FOBT) and then the more invasive structural test is the colonoscopy/sigmoidoscopy (UK colorectal cancer screening pilot group, 2004).

The FOBT is the most commonly used non-invasive CRC screening method and is designed to detect occult (i.e. not accompanied by discernible symptoms or signs) blood in the stool, considering that large polyps and the actual colorectal cancers in the colon and rectum tend to bleed. The test requires the sampler to smear a card with faeces twice after a bowel movement, then repeat this for a further two bowel movements, making a total of six windows covered on the card. The test is then sent in a hygienically sealed envelope back to the relevant clinic and is tested. A positive result is given when any of the windows return as 'positive'. The FOBT has a high false positive rate (2–13%) (Allison *et al.*, 1996) due to the existence of other sources of blood which such test does not take into account, i.e. haemorrhoids and peptic ulcers. The sensitivity (~ 40%) and specificity of

the FOBT are low with respect to detecting early-stage CRC (Niv and Sperber, 1995, Towler *et al.*, 1998). Therefore, most patients who test positive will require an invasive and expensive confirmatory diagnostic procedure such as colonoscopy, which has demonstrated to be a relatively accurate and reliable method for detecting CRC with a sensitivity of approximately 90% (Bretthauer, 2011). Hence, there is a need for a more accurate screening test for colorectal cancer to reduce the number of unnecessary colonoscopies. Being able to screen for colorectal cancer by detecting VOCs present in the headspace of faeces may be a suitable alternative.

1.7.2 Use of volatile biomarkers in screening for colorectal cancer

Normal metabolism generates VOCs that may emanate from faecal matter. Human faecal flora comprises bacteria involved in colonic fermentation producing sulphur containing compounds, such as hydrogen sulphide, dimethyl disulphide, methyl disulphide, and dimethyl trisulphide, and these compounds are responsible for the specific odour of faecal matter (Moore *et al.*, 1987). In the last decade, very little work has been published on the volatiles from stool and their potential in the screening of colorectal cancer. The analysis of VOCs in faecal headspace of neonates showed a large discrepancy in volatiles from premature neonates compared to healthy adults (De Lacy Costello *et al.*, 2008). Indeed, this finding reflects the simplicity of neonatal flora compared to the microbiota of adult gut indicating that most of these VOCs are produced by fermentation of dietary substrates by gut microbes. This study indicates that analysis of VOCs in faecal headspace has the potential for the diagnosis of a range of gastro intestinal diseases, non-invasively and painlessly. However, further studies are necessary to support this experiment and to determine whether analysis of VOCs in faecal headspace can be used to screen for colorectal cancer.

Not only faecal headspace analysis provides insight into the disease but also breath analysis, in which Altomare and co-workers used GC-MS to analyse breath samples from patients with colorectal cancer. The authors established that the pattern of VOCs in patients suffering from colorectal cancer were different from that in healthy controls, particularly levels of some specific VOCs such as 1,3-dimethylbenzene, 1,2-pentadiene, cyclohexene and methylcyclohexene (Altomare *et al.*, 2013), although the confidence of the results is questioned since these VOCs might arise from environmental exposure and the biochemical pathways within the body are unknown.

Chapter 2

Analytical methods for analysing VOCs

The use of VOCs has been discussed previously in disease diagnosis and monitoring and some applications. Detailed information on theory and practice of the instrumental methods for analysing VOCs is now explained below.

2. 1 Gas-phase ion chemistry

Gas-phase ion chemistry is a broad field that has many applications and which embraces several branches of chemistry and physics. It is the science that studies ions and molecules in the gas-phase, most often enabled by some mass spectrometry technique. Gas-phase reactions mainly provide thermochemical and kinetic data which contribute to the fundamental understanding of the reaction mechanisms of a wide range of species from simple inorganic ions to highly reactive organic ions.

Ion-molecule reactions in the gas-phase were first studied under equilibrium conditions in a high-pressure mass spectrometer in 1963 by Kebarle and Godbole for an ion clustering reaction comprising the gas-phase solvation of H_3O^+ (Kebarle and Godbole, 1963). Ion-molecule reactions in the gas-phase proceed via proton transfer, charge transfer, hydride ion transfer, adduct formation, association, oxidation and hydrogen atom transfer (Hoffman and Stroobant, 2007).

A number of instruments and experimental methods have been developed to study these reactions in more detail under highly controlled conditions, such as Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) and Proton Transfer Reaction Mass Spectrometry (PTR-MS), Fourier Transform Mass Spectrometry (FT-MS) and Quadrupole Ion Trap Mass Spectrometry (QIT-MS). Advances in these detection techniques have coincided with the development of new ionization methods such as chemical ionization (CI), fast atom bombardment (FAB) and electrospray (ESI). Together these tools have allowed gas-phase

ion chemistry to be exploited as an analytical tool to monitor chemical species directly, e.g. VOCs in breath.

2. 2 Techniques for VOC measurements

VOC research relies on analytical methods (Table 2.1) that offer high sensitivity, precision and resolution. The on-line, real-time mass spectrometric (MS) techniques SIFT-MS and PTR-MS have been developed with limits of detection ranging from ppbv to pptv, making them ideally suited to VOC analysis (Smith, 2014). Proton transfer reactions occur in both techniques in a chemical ionization process that allows efficient ionization for many organic compounds in the gas phase. Product ion generation in SIFT-MS and PTR-MS is managed using chemical ionization, arising from ion-molecule reactions rather than electron impact or photoionization, leading to much less fragmentation of the molecules. Thus, these techniques are called *soft* ionisation techniques.

SIFT-MS and PTR-MS are well suited to direct, real-time MS profiling without pre-concentration. The MS data sets are produced by either technique are quite simple, easy to handle, including numerous variables perfectly suited for multivariate statistics. Hence both SIFT-MS and PTR-MS are well suited to the work presented in this thesis enabling unequivocal real-time quantification of VOCs, with high sensitivity, precision and resolution.

Gas Chromatography-Mass Spectrometry (GC-MS) combined with thermal desorption (TD) or solid-phase micro extraction (SPME) is also widely used in the analysis of volatile samples. In contrast to SIFT-MS and PTR-MS a major advantage of chromatographic methods are their very high sensitivity due to sample concentration via thermal

desorption (TD) or solid-phase micro extraction (SPME). In addition, the existence of extensive compounds libraries makes compound identification much easier than in SIFT-MS and PTR-MS. In addition, the chromatography component allows compound separation so compound identification of spectra is easier. In this thesis, all the samples were analysed using a GC-MS combined with a thermal desorption unit (TD-GC-MS).

Table 2. 1 – A comparison of the characteristics of the available research techniques in VOC measurements (Adapted from Lourenço and Turner, 2014)

Analytical Method	***Mode of operation***	***Limit of detection (LOD)***	***Sensitivity***	***Specificity***
SIFT-MS	Direct/Real time	ppbv	High	High
PTR-MS	Direct/Real time	pptv	High	Medium-High
IMS	Real-time	ppbv	Medium	Medium
Sensor arrays	Reference to a database	ppbv	Medium	Medium
TD-GC-MS	Pre-concentration	pptv-ppbv	Very-high	Very-high
GC-FID	Direct	ppbv-ppmv	High	High
LAS	Real-time	ppbv	High	High

Other analytical techniques that may be used, include laser absorption spectroscopy (LAS), ion mobility spectrometry (IMS), and a variety of gas sensors and semiconductor-based sensor arrays. Many applications have combined various sensors and materials into a single array, leading to the development of an “electronic nose” (Shurmer *et al.*, 1987).

For the LAS technique, the amount of light absorbed by a sample is related to the concentration of the target specie in the sample. The LAS technique enables quantification of volatiles in exhaled breath down to below ppbv levels. It is particularly useful for monitoring purposes, and, recently, the exhaled breath of healthy volunteers was assessed by LAS-based technique cavity ringdown spectroscopy (CRDS) (Vaittinen *et al.*, 2013).

In the IMS technique, the ions are generated by any of a number of methods (e.g. radioactive strip, electrospray) and they are separated according to their mobilities

through the gas, which is usually air at atmospheric pressure (Borsdorf and Eiceman, 2006) in an applied electric field. During the last ten years the IMS technique has been applied in medical research, such as detection of skin volatiles (Ruzsanyi *et al.*, 2012), detection of volatiles in exhaled breath of patients with lung cancer (Westhoff *et al.*, 2009), and determination of anaesthetics concentration in exhaled breath (Perl *et al.*, 2009).

Sensor technology has been used for many years in clinical testing. However gas sensors are often much less sensitive, usually lack specificity, and are prone to drift. The combination of specificity, selectivity, robustness in operation, reproducible manufacturing uniformity, and long-life stability is not yet offered by current sensors at an acceptable cost level. There are also difficulties in inter-device reproducibility.

2. 3 Mass Spectrometry and Gas-phase ion chemistry

Gas-phase ions are generated and separated according to their mass-to-charge ratio (*m/z*) using electric fields/ magnetic fields. Figure 2.1 shows a schematic representation of how a mass spectrometer works.

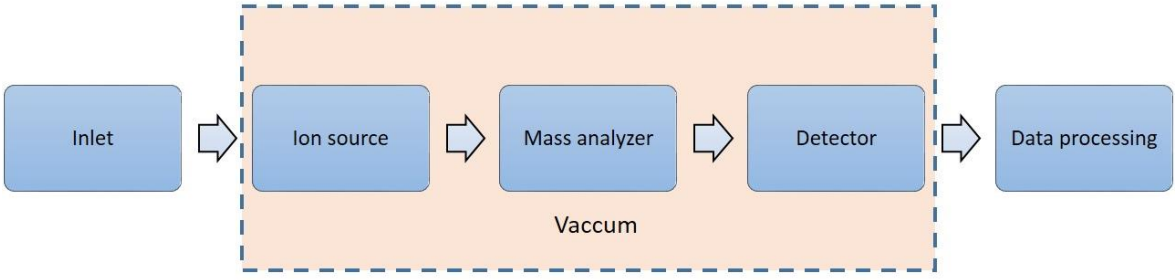


Figure 2. 1 – Components of a mass spectrometer

Overall a mass spectrometer consists of an inlet, an ion source, a mass analyzer, a detector, a high-vacuum system and an acquisition system. The procedure for sample

injection will strongly depend on the physical properties of the sample and it may be accomplished by direct insertion, liquid-chromatography or gas chromatography. The ions are generated in the ion source by many different ways, e.g. electron ionization (EI) or chemical ionization (CI). Current mass analyzers include quadrupole (Q), time-of-flight (TOF), ion trap (IT), magnetic sector field (B), electric sector (E), fourier transform ion cyclotron resonance (FT-ICR), fourier transform orbitrap (FT-OT), in addition to tandem mass spectrometry commonly known as MS/MS, consisting of two mass spectrometers in series connected by a chamber known as collision cell. The ions are recorded by the detector, e.g. electron multiplier, Faraday cup, among others, and the data generated being acquired in the form of a mass spectrum.

2.3.1 Ionization methods

Chemical ionization (CI) uses a molecular ion to transfer charge onto the target analyte (i.e. charge transfer) or through formation and/or the breaking of chemical bonds, such as proton transfer, hydride ion transfer, or ion-molecule adduct formation (Ellis and Mayhew, 2014). Ion-molecule reactions such as proton transfer are often very fast and occur on every collision, which makes them useful for efficient ionization of target molecules in mass spectrometry. Chemical ionization is called a *soft* ionization technique, transferring less energy to the analyte, resulting in less fragmentation of the parent ions.

Electron ionization (EI) can be used for analysing a wide range of VOCs up to 1000 Da molecular weight including hydrocarbons, oils, alkaloids, steroids, pesticides, dioxins, flavours, fragrances, *etc.* EI is often used in conjunction with gas chromatography. In principle, direct analysis of VOCs can be achieved by simply introducing a sample into a commonly used EI ion source of a conventional mass spectrometer. However, the resulting mass spectra obtained by this procedure are extremely complex, making it

difficult to identify a peak that can be attributed to the molecular parent ion. This is due to the high energies employed (~70 eV) which largely exceed the first ionization energies of all chemical compounds. Thus, this excess of energy available generally leads to fragmentation of the molecular ion (Ellis and Mayhew, 2014).

Figure 2.2 shows an EI source. Within the heated ion source operated under high vacuum (10^{-5} to 10^{-6} mbar), electrons emitted by a heated tungsten filament are accelerated to the ionization chamber by a potential difference, usually set to 70 eV. In addition, the electron beam is collimated by a moderate magnetic field generated by two small magnets. The ionization is achieved after collision of the gas sample with the electron beam followed by the removal of an electron from the analyte molecule producing a positively charged ion respectively (Ashcroft and Barnett, 1997).

Equation 2.1

$$e^- + M \rightarrow M^+ + 2e^-$$

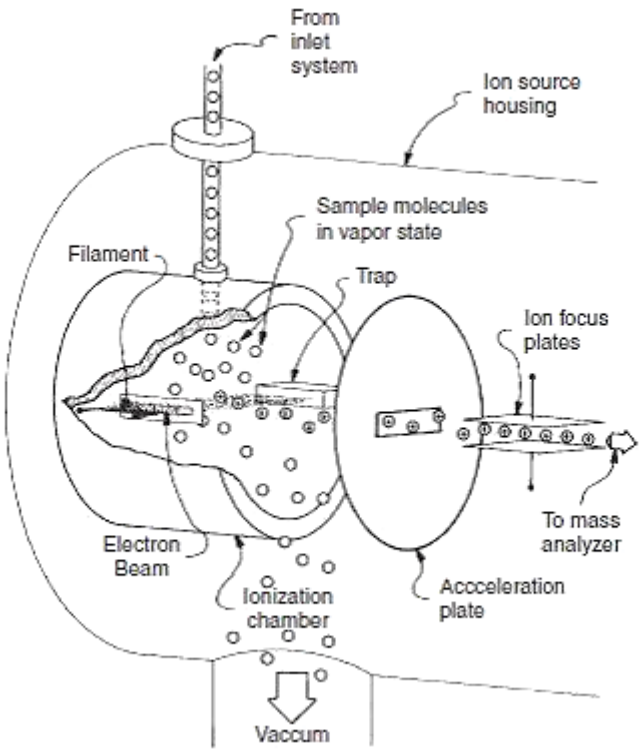


Figure 2. 2 – Electron ionization (EI) source showing the generation of positively charged ions towards the mass analyzer (Sparkman et al., 2011)

Since the internal covalent bonds in the molecule are susceptible to cleavage, most molecular ions tend to fragment due to an excess of energy available, and these fragment ions can in turn fragment leading to a complex product ion pattern. Thus, data analysis becomes complex and difficult to interpret, which delays biomarker discovery accurately.

2.3.2 Gas-phase ion-molecule reactions

Gas-phase ion-molecule reactions yields information on the reactivity of cations or anions in the absence of solvation and counter ion effects. As an illustrative example, Figure 2.3 correlates the energy characteristics of both gas-phase and aqueous solution phase required for the substitution reaction below (Hoffman and Stroobant, 2007).

Equation 2.2

$$Cl^- + CH_3Br \rightarrow CH_3Cl + Br^-$$

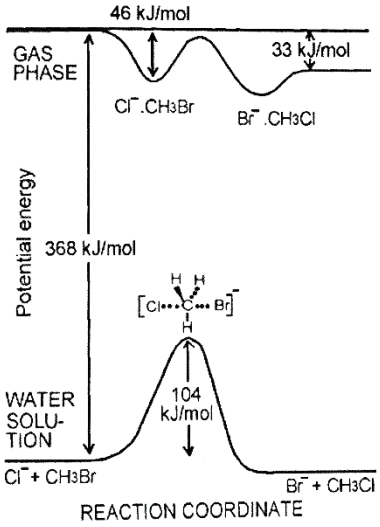


Figure 2. 3 – Potential energy diagram for a substitution reaction in the gas-phase and in solution in water (Hoffman and Stroobant, 2007)

Ion chemistry in the condensed phase is often dominated by ion pairing and solvation interactions which can mask the intrinsic reactivity of the reaction partners. The molecules are strongly solvated, which in turn requires more energy to overcome such a strong activation barrier to form the activated complex. In contrast, chemical reactions in the gas-phase are exothermic, and the activated complex is generated without a

pronounced activation barrier to supress. However, some activation energy is necessary to transform the $Cl^-.CH_3Br$ complex into the activated complex and finally into $Br^-.CH_3Cl$, and the energy produced in the second step is generally lower than the energy produced by the first step. Overall, the actual activation energy in the gas-phase is negative and due to the inexistence of solvation interactions, gas-phase rate coefficients are rather faster than the respective rates coefficients in solution (Hoffman and Stroobant, 2007).

2.3.3 Types of ion-molecule reactions

Ion-molecule reactions in the gas-phase such as those in Table 2.2 are often very fast, which means that they often occur at each collision considering an exergonic reaction where the change in the free energy is negative, indicating a spontaneous reaction (Ellis and Mayhew, 2014).

Table 2. 2 – Ion-molecule reactions in the gas phase.

Ion-molecule reaction	***Designation***
$X^+ + M \rightarrow M^+ + X$	Charge transfer
$XH^+ + M \rightarrow MH^+ + X$	Proton transfer
$X^+ + MH \rightarrow M^+ + HX$	Hydride ion transfer (H^-)
$X^+ + M + Z \rightarrow (M+X)^+ + Z$	Adduct formation
$X^+ + M \rightarrow MX^+$	Association or clustering
$X^+ + MO \rightarrow XO^+ + M$	Oxidation reaction
$X^+ + MH \rightarrow XH^+ + M$	Hydrogen atom transfer

The most abundant types of ion-molecule reactions (Hoffman and Stroobant, 2007) are *charge transfer* reactions. In this process the transfer of charge from the ion to the neutral species occurs. In the case of *proton transfer*, the proton affinity of the proton acceptor (neutral or anion) has to be higher than the proton affinity of the donor (cation or neutral). An *adduct* is formed by direct combination of a neutral species and an "ionizing" ion other than the proton. *Hydride (H^-) ion transfer* occurs every time the H^- ion

is extracted from the carbon group of the neutral molecule to form a different neutral species. In *association* or *clustering* reactions, clusters or aggregates $(n\text{M+H})^+$ are generally observed and result from the combination via noncovalent forces of two or more atoms or molecules of one or more chemical species with an ion. Generally, the proton can be replaced by another cation. The formation of such aggregates is exothermic and its stability is enhanced whether the partners of the aggregate have similar basicities or affinities for the cation, i.e. two or more identical molecules. Commonly, the abundance of the aggregates drops with the increase of the number of associated molecules. *Oxidation* corresponds to the loss of electrons or an increase in oxidation state by a molecule, atom or ion, and usually followed by chemical reduction. *Hydrogen atom transfer* occurs every time a hydrogen free radical is abstracted from the neutral species into the ion.

2.3.4 Proton transfer

Proton transfer ion-molecule reactions in the gas-phase are the basis of both SIFT-MS and PTR-MS and have been used to study the reaction mechanisms of many VOCs. Exothermic proton transfer reactions in the gas-phase tend to be fast ($k \geq 10^{-9}$ cm^3 s^{-1}) and usually proceed at the collisional rate (Blake *et al.*, 2009, Bohme *et al.*, 1980).

<u>Precursor ion formation</u>

H_3O^+, NO^+ and O_2^+ are the most suitable reagent ions used in SIFT-MS analyses of air and breath, because they do not react rapidly with the major components of air and breath, N_2, O_2, H_2O, CO_2 and Ar, but they do react rapidly with most VOCs to produce product ions characteristic of the various neutral molecules, M, in a gas sample being analysed (Smith and Španěl, 1996b). The great majority of PTR-MS experiments carried out to date

have employed H_3O^+ as precursor ion generated by a hollow cathode discharge ion source. The latest instruments use a switchable reagent ion capability, alternating between the three precursor ions, H_3O^+, NO^+, and O_2^+ (Jordan *et al.*, 2009a).

The reaction with water molecules yields the H_2O^+ ions, which in turn are readily converted into H_3O^+ by the fast reaction of direct dissociative ionisation below.

Equation 2.3 $$OH^+ + H_2O \rightarrow H_3O^+ + O$$

Equation 2.4 $$e^- + H_2O \rightarrow H_2O^+ + 2e^-$$

Equation 2.5 $$H_2O^+ + H_2O \rightarrow H_3O^+ + OH$$

In highly humid conditions, the H_3O^+ water cluster ions, $H_3O^+(H_2O)_n$, n = 1,2,3, appear in the product ion mass spectrum (Španěl and Smith, 2000a). Similarly, either for NO^+ or O_2^+, the hydrated cluster ions are produced respectively, $NO^+(H_2O)_n$, n = 1,2 and $O_2^+(H_2O)_n$, n = 1,2 (Španěl and Smith, 2009). These cluster ions can also act as precursor ions and lead to hydrated product ions $MH^+(H_2O)_n$, n = 1, 2, 3.

NO^+ ions formed in the discharge ion source are due to the reactions of O^+ and N^+ atomic ions generated in the ion source from N_2 and O_2 molecules (Španěl *et al.*, 2007b).

Equation 2.6 $$N^+ + H_2O \rightarrow H_2O^+ + N$$

Equation 2.7 $$O^+ + H_2O \rightarrow H_2O^+ + O$$

Equation 2.8 $$O^+ + N_2 \rightarrow NO^+ + N$$

Equation 2.9 $$N^+ + O_2 \rightarrow NO^+ + O$$

O_2^+ ions are formed via direct ionisation of O_2 molecules and via the reaction of O^+ and H_2O^+ ions with O_2 molecules (Španěl *et al.*, 2007b).

Equation 2.10 $$O^+ + O_2 \rightarrow {O_2}^+ + O$$

Equation 2.11 $$H_2O^+ + O_2 \rightarrow {O_2}^+ + H_2O$$

The humidity of the air source gas and the total pressure, determines the total ion currents and the relative signals of the three precursor ion species.

H_3O^+ reactions

Reactions of H_3O^+ ions with VOCs proceed via exothermic proton transfer. Chemical ionization is managed by the proton affinity of the volatile, thus VOCs with proton affinities (PA) higher than proton affinity of water, PA(VOC) > PA(H_2O) = 691 kJ mol^{-1} (Hunter and Lias, 1998), will react and protonated ions (MH^+) will be detected. If this condition is fulfilled, then a reaction occurs in every ion-molecule collision. Reactions of H_3O^+ ions with VOCs occur either in the flow tube of SIFT-MS or drift tube of PTR-MS.

Equation 2.12 $$H_3O^+ + M \rightarrow MH^+ + H_2O$$

Depending on the type of organic molecule under study (i.e. alcohols, aldehydes, ketones, amines, carboxylic acids, *etc.*), MH^+ product ions may dissociate to form ions resulting from the elimination of H_2O molecules.

Equation 2.13 $$H_3O^+ + ROH \rightarrow R^+ + H_2O + H_2O$$

This usually occurs following the protonation of alcohols, aldehydes and carboxylic acids because of the OH group present in these molecules (Španěl and Smith, 1997, Španěl *et al.*, 1997, Španěl and Smith, 1998c).

The ether reactions are quite complex. In this situation, following the protonation of the ether, MH^+ ions are usually generated from the elimination of alcohol and hydrocarbon molecules (Španěl and Smith, 1998d). Elimination of alcohol is also common in the esters reactions (Španěl and Smith, 1998c). The equations below illustrate the process in ethers, R being a radical such as CH_3.

Equation 2.14
$$H_3O^+ + R_xOR_y \rightarrow R_xOR_yH^+ + H_2O$$

Equation 2.15
$$H_3O^+ + R_xOR_y \rightarrow {R_x}^+ + R_yOH + H_2O$$

Equation 2.16
$$H_3O^+ + R_xOR_y \rightarrow {R_yOH_2}^+ + R + H_2O$$

Previous studies in amines reactions showed that they are efficiently protonated by H_3O^+ due to its high proton affinity (Španěl and Smith, 1998a). In the reactions of the secondary amines with H_3O^+ another reaction process is observed, the loss of an H_2 molecule from the excited protonated parent molecule (Španěl and Smith, 1998a).

Equation 2.17
$$H_3O^+ + R_2NH \rightarrow R_2N^+ + H_2 + H_2O$$

Remarkably, the reactions of some aliphatic hydrocarbons and halocarbons with H_3O^+ proceed via ion-molecule association (Španěl and Smith, 1998a).

Equation 2.18

$$H_3O^+ + M + He \rightarrow H_3O^+M + He$$

This unusual reaction apparently occurs when the proton affinity of the VOC is less than the proton affinity of water, PA(VOC) < PA(H_2O) = 691 kJ mol^{-1}. However, the study of lower order hydrocarbons is impossible to carry out, because its proton affinity is excessively smaller than PA(H_2O), thus neither proton transfer nor association occurs (Španěl and Smith, 1998a).

NO^+ reactions

The reactions of NO^+ ions with VOCs usually procced via charge transfer. However, it can be followed by one or two of these processes in parallel, hydride ion transfer, hydroxide ion transfer, alkoxide ion transfer, and ion-molecule association. Most amines proceed via charge transfer with NO^+, and sometimes the process is observed in parallel with hydride (H^-) ion transfer (Španěl and Smith, 1998a). The ionization energy of NO is 9.26 eV (Lias *et al.*, 1988), so it ionizes fewer compounds via the charge transfer mechanism than does O_2^+. For organic molecules satisfying the following conditions, IE(VOC) < IE(NO^+) = 9.26 eV, a reaction via charge transfer is guaranteed.

Equation 2.19

$$NO^+ + M \rightarrow M^+ + NO$$

Hydride (H^-) ion transfer and hydroxide (OH^-) ion transfer are the most common reactions occurring in primary and secondary alcohols with NO^+. However, the reactions of tertiary alcohols proceed only via hydroxide (OH^-) ion transfer (Španěl and Smith, 1997). In ethers and saturated aldehydes, hydride (H^-) ion transfer is the only process that occurs in the reactions with NO^+ (Španěl and Smith, 1998d, Španěl *et al.*, 1997). In hydride ion transfer the H^- ion is extracted from the $–CH_3$ carbon group of either alcohol, ether, or aldehyde.

Equation 2.20

$$NO^+ + ROH \rightarrow [ROH - H]^+ + HNO$$

Equation 2.21

$$NO^+ + ROH \rightarrow [ROH - OH]^+ + HNO_2$$

Ion-molecule association is common in the reactions between NO^+ and carboxylic acids and esters (Španěl and Smith, 1998c), and ketones (Španěl *et al.*, 1997). Association reaction is indicated below.

Equation 2.22

$$NO^+ + M \rightarrow MNO^+$$

In NO^+ reactions with carboxylic acids, association is seen in parallel with hydroxide (OH^-) ion transfer. In the case of ester reactions with NO^+, association often occurs in parallel with alkoxide (OR^-) ion transfer. Alkoxide ion transfer is indicated below.

Equation 2.23

$$NO^+ + RCO_2R \rightarrow NO^+RCO_2R + He$$

Equation 2.24

$$NO^+ + RCO_2R \rightarrow RCO^+ + RNO_2 + He$$

Alkenes proceed via charge transfer due to its small IE < IE(NO^+) = 9.26 eV (Španěl and Smith, 1998a).

O_2^+ reactions

The reactions of O_2^+ ions with VOCs mostly proceed either by charge transfer, where IE(VOC) < IE(O_2^+) = 12.07 eV (Lias *et al.*, 1988), or dissociative charge transfer yielding two or several fragment ions. Such reactions are defined below.

Equation 2.25

$$O_2^+ + M \rightarrow M^+ + O_2$$

Equation 2.26

$$O_2^+ + M \rightarrow Fragment^+ + neutral\ products$$

The efficiency of the charge transfer process is often less than 100%. The reactions of O_2^+ ions are most used for the quantification of small molecules that do not react with either H_3O^+ or NO^+ ions, such as small hydrocarbons. Reactions of O_2^+ ions are less useful because their reactions with polyatomic ions usually result in multiple ion fragments (Smith and Španěl, 2005).

2.3.5 Theoretical prediction of proton transfer rate coefficients

Quantification of VOC concentrations may be accomplished if proton transfer rate coefficients are known. Proton transfer reactions occur efficiently if the proton affinity of the acceptor molecule exceeds the proton affinity of the donor (H_2O), thus such reactions are exothermic, fast ($k \geq 10^{-9}$ cm^3 s^{-1}) and usually proceed at the collisional rate (Blake *et al.*, 2009, Bohme *et al.*, 1980).

The simplest theory to determine rate coefficients of exothermic ion-molecule reactions is given by the calculations using the Langevin (Langevin, 1905) theory. The Langevin theory provides rate coefficients, k_L, between a point of charge (ion) and a non-polar neutral molecule; assuming there is no permanent dipole; and interaction occurs through an induced dipole.

Equation 2.27

$$k_L = \sqrt{\frac{\pi\alpha q^2}{\mu\varepsilon_0}}$$

where, q represents the charge of the ion (1.60×10^{-19} C); α is the polarizability of the neutral molecule (m^3); ε_0 represents the permittivity of free space (8.85×10^{-12} J^{-1} C^2 m^{-1}); μ stands for the reduced mass of colliding partners ($\mu = m_1 \times m_2/(m_1 + m_2)$, where m_1 and m_2 are the masses of the ion and the neutral molecule respectively.

In polar molecules, such as many VOCs, it is necessary to account for the ion-induced dipole interaction between the ion and the permanent dipole of the neutral molecule. This interaction is accounted for in the second term of the following equation developed by Su and Bowers (Su and Bowers, 1973), and known as The Average Dipole Orientation (ADO) theory:

Equation 2.28

$$k_{ADO} = \sqrt{\frac{\pi\alpha q^2}{\mu\varepsilon_0}} + \frac{C\mu_D q}{\varepsilon_0} + \sqrt{\frac{1}{2\pi\mu k_B T}}$$

where, k_B represents the Boltzmann constant (1.38×10^{-23} m^2 kg s^{-2} K^{-1}); T is the temperature of 300 K; μ_D is the dipole moment expressed in Debye (D); C accounts for the average orientation of the permanent dipole moment of the molecule, and is a function of dipole moment (μ_D) and polarizability (α) which can be determined elsewhere (Ellis and Mayhew, 2014).

Su and Chesnavich parameterised trajectory calculations (Su and Chesnavich, 1982) to model the reaction process of thermal ion-polar molecule capture collisions to derive capture rate coefficients, k_{cap}, which in turn have been widely determined and then used in the determination of VOC concentration via SIFT-MS. The rate coefficient, k_{cap}, is given by:

Equation 2.29

$$k_{cap} = k_L \times K_{cap}(T_R, I^*)$$

where, k_L is the Langevin rate coefficient. K_{cap} (T_R, I^*) is a parameterised quantity dependent on two reduced parameters T_R and I^*. Where, $T_R = 4\pi\varepsilon_0 \times (2\alpha k_B T/\mu^2_D)$; $x = \sqrt{(1/T_R)}$; and $I^* = \mu_D I/\alpha q\mu$ (Ellis and Mayhew, 2014).

Considering that K_{cap} is insensitive to the value of I^*, thus:

Equation 2.30

$$K_{cap} = \begin{cases} \dfrac{(x + 0.5090)^2}{10.526} + 0.9754\,, & x \leq 2 \\ 0.4767x + 0.6200\,, & 2 \leq x \leq 3 \\ 0.5781x + 0.3165\,, & 3 \leq x \leq 35 \\ 0.6201x - 1.153\,, & 35 \leq x \leq 60 \\ 0.6347x - 2.029\,, & x \geq 60 \end{cases}$$

The theories discussed so far do not consider the effect of a homogenous electric field on the rate coefficients. Ion-molecule reactions in the drift tube of PTR-MS proceed with energies larger than thermal. This means that ions travel with an additional energy provided by the applied electric field. Thus, it is necessary to consider the centre-of-mass kinetic energy, KE_{cm}, (Lindinger *et al.*, 1998) for a collision between an ion and the neutral molecule:

Equation 2.31

$$KE_{cm} = \frac{m_n}{m_{ion} + m_n}\left(KE_{ion} - \frac{3}{2}k_B T\right) + \frac{3}{2}k_B T$$

where, m_{ion} is the mass of the ion, and the mass of the neutral molecule, m_n. The mean kinetic energy of the ions, KE_{ion}, is expressed by the Wannier expression (Wannier, 1953):

Equation 2.32

$$KE_{ion} = \frac{3}{2}k_B T + \frac{1}{2}m_{ion}v_d^2 + \frac{1}{2}m_b v_d^2$$

where, m_b is the buffer gas molecules (air) and the velocity of the drift tube, v_d, is given by the following equation:

Equation 2.33

$$v_d = \mu_0 \times N_0 \times \frac{E}{N}$$

μ_0 represents the reduced ion mobility for H_3O^+ in air and this can be experimentally determined as a function of *E/N.* This can be achieved by sending a short pulse of ions along the drift tube and measuring the time of arrival at the end. These values are reported elsewhere (Ellis and Mayhew, 2014) and in this work an average value of 2.8 cm^2 s^{-1} V^{-1} has been used; N_0 stands for the gas number density under standard conditions of temperature and pressure (273.15 K and 760 Torr = 1.01325 × 10^5 Pa); *E/N* ratio, where *E* is the electric field strength and *N* is gas number density in the reaction chamber. The drift velocity increases either by increasing the electric field strength on PTR-MS, or by decreasing the pressure or temperature of the drift tube.

The effective temperature, T_{eff}, describes the internal energy of ion-neutral molecule collisions and depends only on the ratio *E/N* (Viggiano and Morris, 1996). The following expression can be used to estimate the effective ion translational temperature and may be derived from the equation 2.31 by imposing $KE_{cm} = 3/2k_B T_{eff}$, thus getting:

Equation 2.34

$$T_{eff} = T + \frac{v_d^2}{3k_B}\left[\frac{m_n(m_{ion} + m_b)}{m_{ion} + m_n}\right]$$

Su published a method to calculate collisional rate coefficients, k_{cap} (T_{rot}, KE_{cm}), and parameterising k_{cap} explicitly as a function of the centre-of-mass kinetic energy, KE_{cm}, and not just of the temperature (Su, 1994). Based on trajectory calculated results relating the rate coefficient given by the Langevin theory (k_L) to the capture rate constant (k_{cap}) by a factor, K_c, parameterised by KE_{cm} and rotational temperature, T_{rot}.

Equation 2.35

$$k_{cap}\,(T_{rot}, KE_{cm}) = k_L \times K_c$$

K_c is given by the following equation, where τ and ε are given respectively.

Equation 2.36

$$K_c = 1 + (0.727143 \times \tau^{0.4} \times \varepsilon^2 \times S) + [3.71823 \times (1 - S)] \times \left[\sin\left(0.58692 \times (4.97894 + \ln \tau)\right) \times \tau^{0.6} \times (\varepsilon - 0.5)^{0.5}\right]$$

Depending on two reduced parameters, namely, $\tau = \mu_D / \sqrt{(\alpha T_{rot})}$ and $\varepsilon = \mu_D / \sqrt{(\alpha KE_{cm})}$.

2. 4 SIFT-MS

Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) was conceived and developed by Adams and Smith in 1976 (Adams and Smith, 1976), for the study of ion-neutral reactions at thermal interaction energies (Smith and Adams, 1987). Initially it was developed to satisfy the need of kinetic data on gas-phase ion-neutral reactions observed in cold interstellar clouds (Smith, 1992). In 1996, it became a method focused on real-time, on-line analysis of volatile trace gases of biological origin with medical applications (Španěl *et al.*, 1996, Španěl and Smith, 1996, Smith and Španěl, 1996a, Smith and Španěl, 1996b), such as clinical diagnosis, therapeutic monitoring and physiological studies (Smith and Španěl, 2007, Smith and Španěl, 2005, Španěl *et al.*, 1996, Španěl and Smith, 2007).

SIFT studies are carried out under well-controlled thermalised conditions at 300 K, allowing the determination of kinetic data (i.e. rate coefficients) and these are used to derive the analyte concentrations. Thus, quantification of VOCs in air is achieved by using an in-built kinetics library, although it is good practice to periodically check the quantification using known standards.

With this real-time device (Figure 2.4) the patient may breathe directly into the equipment through a heated capillary in order to avoid condensation in the capillary. Alternatively, headspace samples may also be analysed, i.e. those produced in Nalophan bags containing the VOC or other material suited for this type of experiment, which ideally does not emit background contaminants.

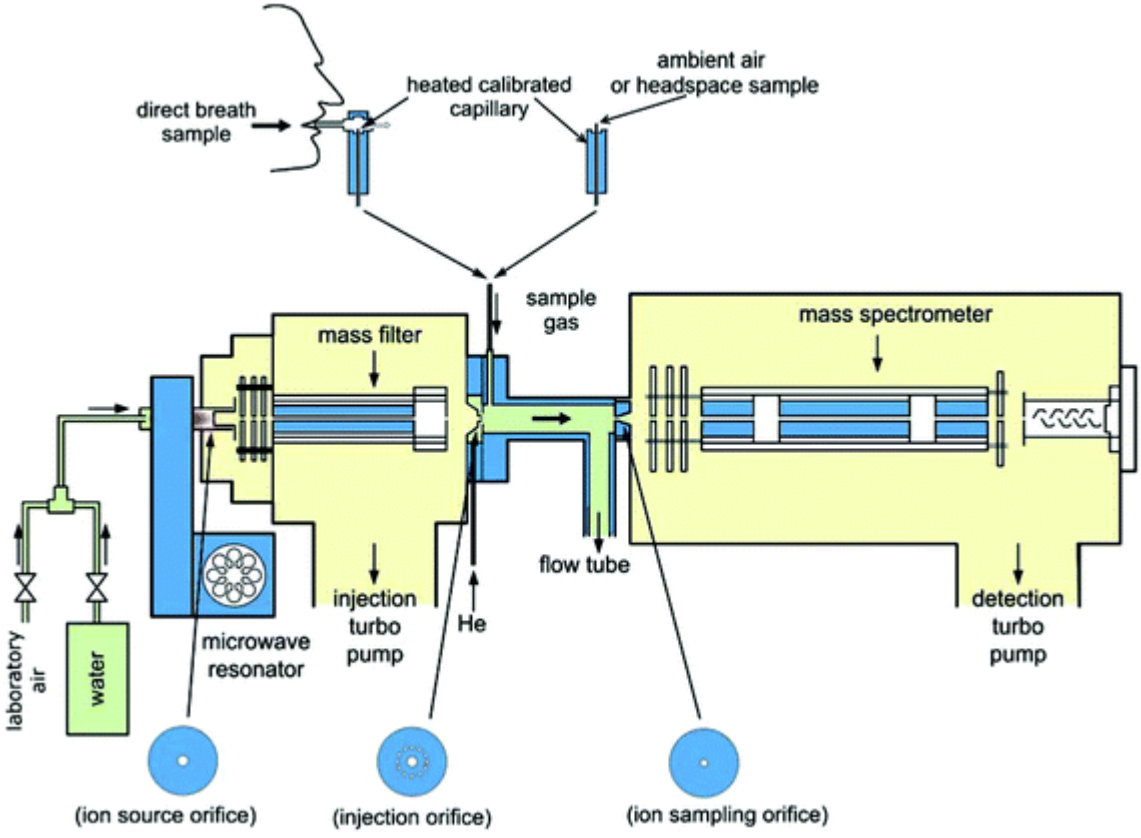


Figure 2. 4 – A schematic of the selected ion flow tube apparatus, SIFT (Smith and Španěl, 2015).

With the SIFT-MS technique, it is possible to identify and differentiate isomers by using the different precursor ions (H_3O^+, NO^+, O_2^+) and by applying two different modes of analysis. The full scan mode (FS) produces a complete mass spectrum, and a multiple ion mode (MIM) monitors the count rates of the selected precursor ion and selected product ions in real-time. A complete mass spectrum is obtained by sweeping the downstream quadrupole mass analyzer over a selected mass-to-charge ratio (*m/z)* range for a selected period of time. The count rates of the ions are then calculated from the numbers of counts and the total sampling time for each ion. In the MIM mode the downstream quadrupole mass analyzer is rapidly switched between the masses of all of the precursor ions and the selected products ions, briefly pausing on each of these masses for a short time period. The partial pressures of the trace gases are calculated as a function of time (Smith and Španěl, 2005).

The precursor ions are produced in an ion source – microwave discharge ion source. These microwave discharges are established in a glass tube separated from the upstream quadrupole mass filter by a disc in which there is a central orifice that is small enough to allow pressures approaching 1 Torr to be established in the ion source whilst ensuring that the pressure in the mass filter is maintained at about 10^{-4} Torr by a turbomolecular pump.

The ions are selected in the upstream quadrupole mass filter according to their mass-to-charge ratio, and injected into a fast-flowing carrier gas (usually pure helium at a pressure of approximately 1 Torr) through a Venturi inlet, in order to minimize backflow of the carrier gas into the mass filter. Both quadrupoles of the instrument require a pressure of $\leq 10^{-4}$ Torr for effective operation, while the flow tube is set at pressures near 0.6 Torr. An excess of carrier gas is used to certify that the reactant ions are formed well

upstream of the point where the neutral reagent gas is added, ensuring ion-molecule reactions at thermal energies.

The chemical reaction takes place in the flow tube, where the selected precursor ion rapidly reacts with the sample gas, a process achieved by soft chemical ionization of the sample. Reactions of H_3O^+ ions with organic compounds proceed via exothermic proton transfer. The reactions of NO^+ ions with organic molecules proceed via charge transfer, hydride transfer, hydroxide transfer, alkoxide transfer, and ion-molecule association. For O_2^+ precursor ions, the ionization process mostly proceed either by charge transfer reaction or dissociative charge transfer (Smith and Španěl, 2005).

In the downstream quadrupole mass analyzer, different count rates of product ions are selected and further detection is achieved in an electron multiplier, for example a *channeltron*.

Satisfactory pressure regimes are obtained throughout the instrument by a two-step vacuum. A vacuum pump delivers the preliminary vacuum necessary for the start-up of the mass filter, followed by the action of a high-speed pump at the downstream end of the flow tube.

Compared to PTR-MS, SIFT-MS is less sensitive due to the existence of a mass filter, which selects the precursor ion to be used according to its mass-to-charge ratio, however, recent developments have led to improved sensitivity in SIFT-MS (Spesyvyi *et al.*, 2015). A uniform electric field has been incorporated along the flow-drift tube to supress ion diffusive loss, thus increasing analytical sensitivity, and to allow a smaller, lower speed

carrier gas drive pump to be used. The kinetic and internal energies of the ions can be altered by changing the *E/N* value.

2.4.1 Quantitative determination of VOC concentration in SIFT-MS

Quantitative determination of VOC via SIFT-MS is achieved by means of using an in-built kinetics library, where *k* is determined through Su and Chesnavich (Su and Chesnavich, 1982) trajectory calculations. Quantification depends on reaction rate coefficient; reaction time; temperature; sample flow rate; flow tube pressure; reagent and product ion intensity; and mass discrimination (diffusion loss).

Equation 2.37

$$[M] = \frac{1}{t} \frac{{I_{p1}}/{D_{ep1}} + {I_{p2}}/{D_{ep2}} + \cdots}{I_{i1}k_1 + {I_{i2}\frac{k_1 + k_2}{2}}/{D_{ei2}} + \cdots}$$

k is the rate coefficient for the reaction and *t* is the reaction time. D_e is a differential diffusion enhancement coefficient that accounts for the fact that the reagent ions and the product ions diffuse through the helium carrier gas to the walls of the SIFT-MS flow tube at different rates. Mass discrimination occurring in the quadrupole mass analyzer, is accounted for to obtain accurate quantitative analyses. I_{p1}, Ip_2, etc. are the count rates (corrected for mass discrimination in the quadrupole mass analyzer) for all the product ions; I_{i1}, I_{i2}, etc. are the count rates (also corrected for mass discrimination) of the precursor ions; k_1 are the rate coefficients for the reactions of the precursor ions with each molecule (M). The rate coefficients k_2, etc. are the rate coefficients for the reactions of the hydrated precursor ions with each M; D_{ep1}, D_{ep2}, etc. are the respective differential diffusion enhancement correction factors (Smith and Španěl, 2005).

2. 5 PTR-MS

Proton transfer reaction mass spectrometry (PTR-MS) was developed in the mid-90s by Lindinger and co-workers (Lindinger *et al.*, 1998). This technique allows real-time, on-line determination of concentrations of VOCs, with detection sensitivity greater than SIFT-MS. The first PTR-MS instruments developed only used H_3O^+ as precursor ion, a disadvantage compared to SIFT-MS. However, the latest instruments use a switchable reagent ion capability like SIFT-MS (Jordan *et al.*, 2009a). There are usually overlapping ions in clinical sample headspace and the use of PTR-MS equipped with a time-of-flight (TOF) mass analyser improves mass resolution to assist ion identification (Herbig *et al.*, 2009, Jordan *et al.*, 2009b).

PTR-MS employs an electric field, *E*, along the drift tube axis to increase the velocities of the ions. The change of the ratio *E/N*, where *N* is the number density of the drift tube buffer gas molecules, will affect the reagent ion hydration and product ion fragmentation. Under normal operating conditions *E/N* is in the range 120–130 Td, representing a compromise between reagent ion hydration on the one hand, and molecular (product) ion fragmentation on the other. The electric field also prevents the formation of substantial quantities of cluster ions. In contrast to SIFT-MS, PTR-MS operates at higher effective temperatures and thus the underlying ion chemistry is often not known (Beauchamp *et al.*, 2013, Hewitt *et al.*, 2003). The recent advances in PTR-MS technology have demonstrated a diverse range of applications, especially for breath gas analysis (Herbig and Amann, 2009).

A PTR-MS instrument is often much shorter than a SIFT-MS, due to its shorter drift tube, has a typical length being 10 – 20 cm, and, consequently, the pumping system is reduced

in size, making PTR-MS suitable for transport and field studies. However, the latest advances in SIFT-MS have surpassed this issue, in which the instrument has become much smaller and suitable for transport (Španěl and Smith, 2013).

The measurements discussed in chapter 4 were acquired with a commercial Series I PTR-TOF-MS (Kore Technology Ltd.) (Figure 2.5) equipped with an ion funnel, although for this study the ion funnel facility was not used. Note that in the schematic representation indicated in Figure 2.5 the drift tube is not equipped with an ion funnel, which has been added later on. The instrument delivers: mass resolution ≈ 2000-2500 (FWHM), and sensitivity for benzene > 200 cps/ppbv. This instrument uses a *reflectron* TOF-MS, another way to improve mass resolution, in which the trajectories of all of the ions are deliberately reversed part way along their journey to the detector, as illustrated in Figure 2.5. The hollow cathode discharge ion source (Figure 2.6) used in PTR-MS is maintained at about 1 mbar and consists of two planar electrodes able to produce an electrical discharge in a gas. A high voltage is established between the two electrodes, cathode negatively charged and anode positively charged. Neutral molecules are ionised by collisions with high energy electrons to produce primary ions (Ellis and Mayhew, 2014).

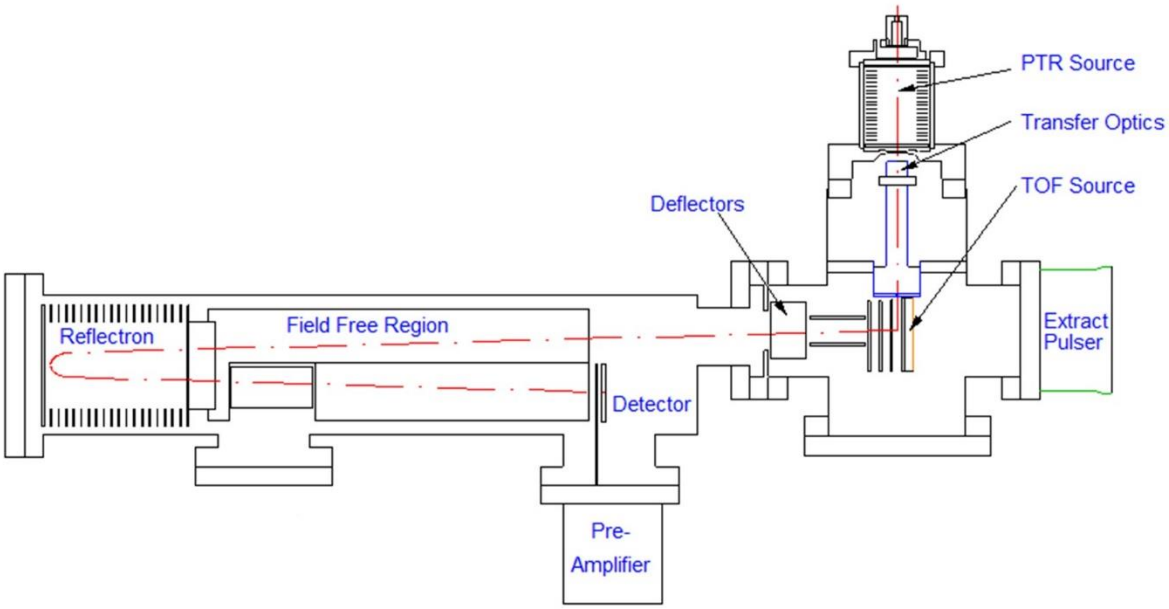


Figure 2. 5 – A schematic of the Series I KORE PTR-MS used for this work equipped with a reflectron TOF-MS. Note that in this schematic the drift tube is not equipped with an ion funnel (Adapted from Kore Technology).

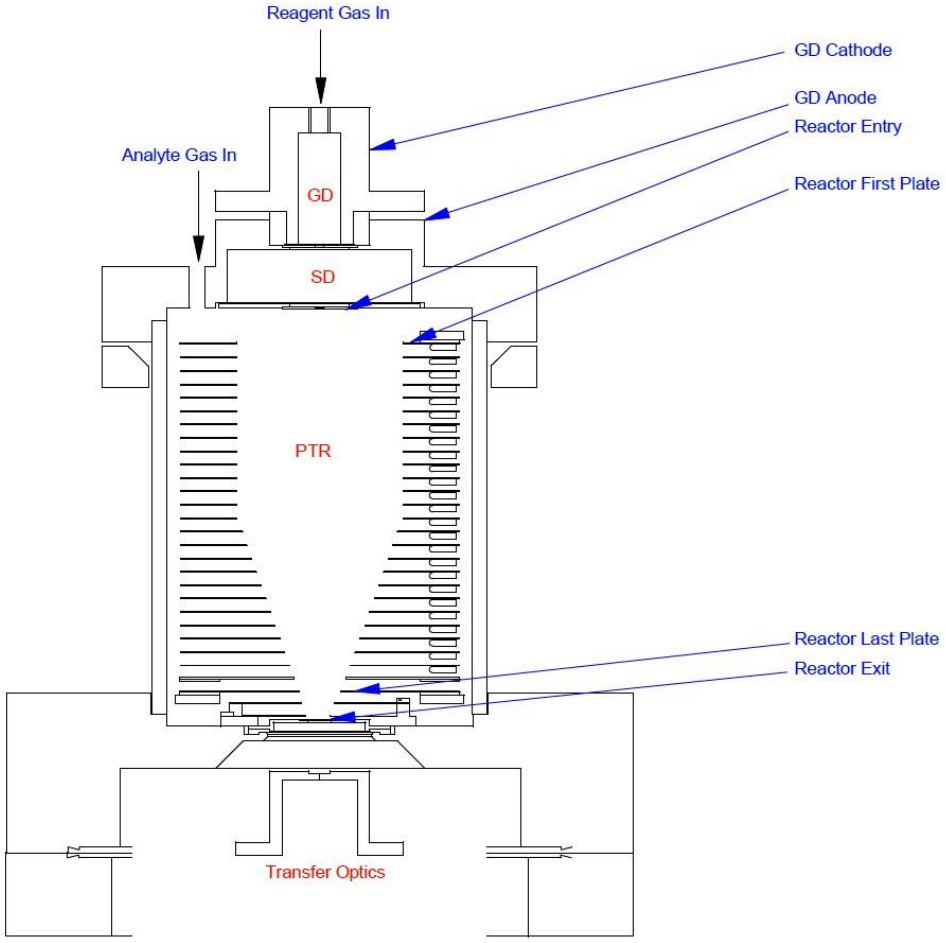


Figure 2. 6 – A schematic of the Series I KORE PTR-TOF-MS used for this work. Features of the PTR ion source adapted with a hollow cathode discharge ion source containing two electrodes, one cathode negatively charged and anode positively charged (Adapted from Kore Technology).

Other mass analyzers such as quadrupoles create spectra by scanning the allowed mass through the mass-range of interest. This is the fundamental reason why a TOF-MS can have a very high overall sensitivity compared to other mass analyzers. Not having to scan in the conventional sense also means that TOF mass analyzers have no difficulties with high mass ions; it is just a matter of waiting for them to arrive.

The growing need for analytical methods with great sensitivity in the analysis of VOCs led to the development of the PTR-QiTOF (Sulzer *et al.*, 2014), whereas "Qi" stands for "Quadrupole interface". In contrast to commercially available PTR-TOF-MS devices so far, which utilize a transfer lens system, the novel prototype is equipped with a quadrupole ion guide for the highly effective transfer of ions from the drift tube to the mass spectrometer. This innovation in the PTR-MS system greatly improves the TOF mass resolution, approximately 10400, because of the encouraging injection conditions.

Like all chemical ionisation techniques, PTR-MS cannot separate isomers, which usually creates difficulties in the qualitative analysis of VOCs. To overcome this downside of PTR-MS, gas chromatography has been coupled with PTR-MS, but this will greatly reduce the time resolution due to the relatively slow transit times of compounds through a GC column (Fall *et al.*, 2001, Lindinger *et al.*, 2005). Recent development of fast gas chromatography (fastGC) coupled with PTR-MS promises much faster separation (Romano *et al.*, 2014, Materić *et al.*, 2015). The capability of fast temperature ramps and short/ thin-film capillary columns combined with a fast and sensitive detector may provide resolution times of less than 5 minutes.

2.5.1 Theoretical determination of VOC concentration in PTR-MS

In contrast to SIFT-MS in PTR-MS the underlying ion chemistry is often not known, specifically, the kinetics of the ion-molecule reactions and reaction time are not well established and can be very sensitive to changes in *E/N*. Thus, careful calibration of the instrument is usually carried out for each VOC. Nevertheless, quantification of VOC concentrations may be accomplished if the reaction time and proton transfer rate coefficients are known (Beauchamp *et al.*, 2013). As PTR-MS is a real time soft ionisation technique, such protonated parent ions usually are the main product ions making identification much easier than with other analytical techniques. Quantification is directly dependent on the proton transfer rate coefficient, therefore it is essential to stress the importance of the gas-phase ion chemistry studies on ion-molecule reactions. Theoretical determination of VOC concentration via PTR-MS is discussed in chapter 4, where standards calibrated for the gas-phase were generated according to Henry´s law and theoretical prediction of reaction rate coefficients were determined, and further calculation of VOC concentrations derived.

VOC concentration in ppbv is determined according to the following expression:

Equation 2.38

$$[R] = \frac{[R_{ppb}]}{[air]}$$

where, *R* is VOC concentration in ppb by volume of air, i.e. ppbv, and the density of air in the reaction chamber is given by:

Equation 2.39

$$[air] = \frac{273.15}{T_d\,[K]} \cdot \frac{N_A}{22400} \cdot \frac{p_d\,[mbar]}{1013\,[mbar]}$$

where, p_d represents the pressure and T_d the temperature in the drift tube, in mbar and K respectively. 22400 indicates conversion from mole-volume using Avogadro's number $N_A = 6.022 \times 10^{23}$ cm^{-3}.

Gas-phase ion-molecule reactions occur in the drift tube, according to:

Equation 2.40
$$H_3O^+ + R \xrightarrow{k} RH^+ + H_2O$$

A first approximation to determine [R] was given by Lindinger, considering [RH^+] << [H_3O^+] ≈ [H_3O^+]$_0$ = constant (Lindinger *et al.*, 1998):

Equation 2.41
$$[RH^+] = [H_3O^+]_0\left(1 - e^{-k[R]t}\right) \approx [H_3O^+]_0[R]kt$$

Theoretical determinations of *R* expressed in ppbv (Beauchamp *et al.*, 2013), i.e. the VOC concentration in PTR-MS, may be accomplished via the following equation (equation 2.42):

Equation 2.42
$$[R] = \frac{1}{kt} \cdot \frac{i(RH^+)_{cps}}{i(H_3O^+)_{cps}} \cdot \frac{T_d}{T_0} \cdot \frac{22400}{N_A} \cdot \frac{p_0}{p_d} \cdot \frac{Tr_{H3O+}}{Tr_{RH+}}$$

where, *k* represents the proton transfer rate coefficient expressed in cm^3 s^{-1} between the hydronium ion (H_3O^+) and R; *t* expressed in seconds is the reaction time of the ions in the drift tube and is usually fast, typically 100 μs, where $t = L^2/\mu U$ and $\mu_0 = \mu\, p_d\, T_0/T_d\, p_0$. *L* is the drift tube length in cm; U is the electric potential applied to the drift tube in Volts; μ expressed in cm^2 s^{-1} V^{-1} is the mobility of the H_3O^+ in the drift tube which depend on the working conditions (pressure and temperature). μ_0 represents the reduced ion mobility for H_3O^+ in air and for this work it has been considered an average value of 2.8 cm^2 s^{-1} V^{-1}.

$i(RH^+)_{cps}$ stands for the counts per second of detected protonated parent ion and its fragments; $i(H_3O^+)_{cps}$ is the counts per second of hydronium ion. For highly humid samples, a better approximation would consider the presence of H_3O^+ water clusters $[H_3O(H_2O)_n]^+$ ions, with n = 0, 1, 2 , 3 (the extension of which depends upon reduced electric applied). T_d and T_0 represent the temperature (K) of the drift tube and standard temperature (273.15 K), respectively; 22400 indicates conversion from mole-volume using Avogadro's number N_A = 6.022 × 10^{23} cm^{-3}; p_0 and p_d represent the standard pressure (1013.25 mbar) and the pressure in the drift tube in mbar respectively. *Tr* terms indicate the mass-dependent transmission efficiency of ions into the PTR-TOF-MS detection system, where $Tr_{H3O+} = \sqrt{(m/z)_{H3O+}}$ and $Tr_{RH+} = \sqrt{(m/z)_{RH+}}$ (Cappellin *et al.*, 2012).

Depending on the working conditions, the protonated parent ion might undergo fragmentation and the contribution of all fragment ions must be taken into account (equation 2.43):

Equation 2.43

$$[R]_{total} = [R] + \sum_i \frac{1}{kt} \cdot \frac{i(RH_i^+)_{cps}}{i(H_3O^+)_{cps}} \cdot \frac{T_d}{T_0} \cdot \frac{22400}{N_A} \cdot \frac{p_0}{p_d} \cdot \frac{Tr_{H3O+}}{Tr_{RH_i^+}}$$

2. 6 Gas chromatography-mass spectrometry combined with thermal desorption

Gas chromatography coupled with a mass spectrometer allows qualitative and quantitative analysis of complex mixtures of VOCs at trace concentration. This hyphenated technique uses a mass spectrometer as the GC detector to identify and quantify unknown samples. Isomers are therefore easier to identify due to the separation of compounds achieved in the column, an advantage compared to the common mass spectrometric techniques which usually deliver overlapping peaks in clinical samples.

Usually, the concentration of the VOC of interest is too low for the direct measurement of an air sample, and therefore enrichment on suitable adsorbents is necessary.

Gas chromatography-mass spectrometry (GC-MS) combined with thermal desorption (TD) (Figure 2.7), or solid-phase micro extraction (SPME) has been widely used for trace gas analysis though it has some limitations. For example, pre-concentration of the sample is required prior to analysis through pumping significant volumes of trace gases samples onto adsorption traps; VOCs with low molecular weight can be difficult to accurately quantify; it is time consuming; it is not a real-time monitoring technique; GC-MS requires calibration for each trace gas; and some compounds are separated better than others. Thermal desorption (TD) relies on the use of heat, up to 350 °C, to directly extract analytes from sample and swept onto GC column, with no chemical bonds broken. Nevertheless, TD-GC-MS was used to test all the volatile samples analysed in this thesis.

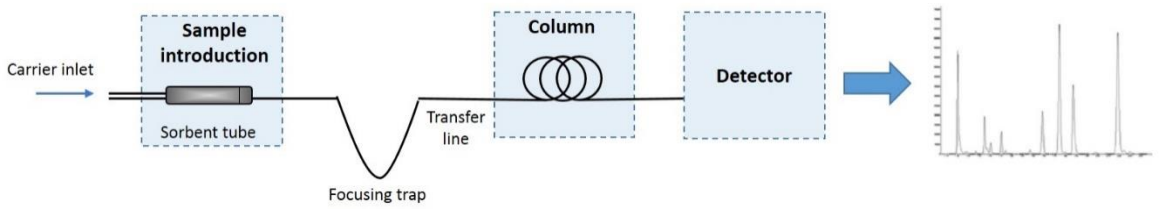


Figure 2. 7 – A schematic representation of a thermal desorption GC-MS system (Adapted from Poole, 2012).

Electron ionization is often used in a GC-MS system, producing positively charged ions, in conjunction with a mass analyzer (e.g. quadrupole or TOF) allowing separation of compounds according to their mass-to-charge (*m/z*) ratio and further detected. If electron ionization is being used, a spectrum of the analyte can usually be found in the *NIST Mass Spectral Database*.

2.6.1 Calibration with internal standards

Thermal desorption procedures are generally calibrated using internal standard methods, such as internal standard addition using deuterated toluene or bromofluorobenzene. Calibration is accomplished by loading standards onto sample tubes and desorbing them through the two-stage thermal desorption process. Standard solutions are prepared and small volumes (typically 0.2 - 2 µL) loaded into each TD tube individually (Poole, 2012). Markes' Calibration Solution Loading Rig (CSLR™) (Figure 2.8) was specifically designed for this purpose.

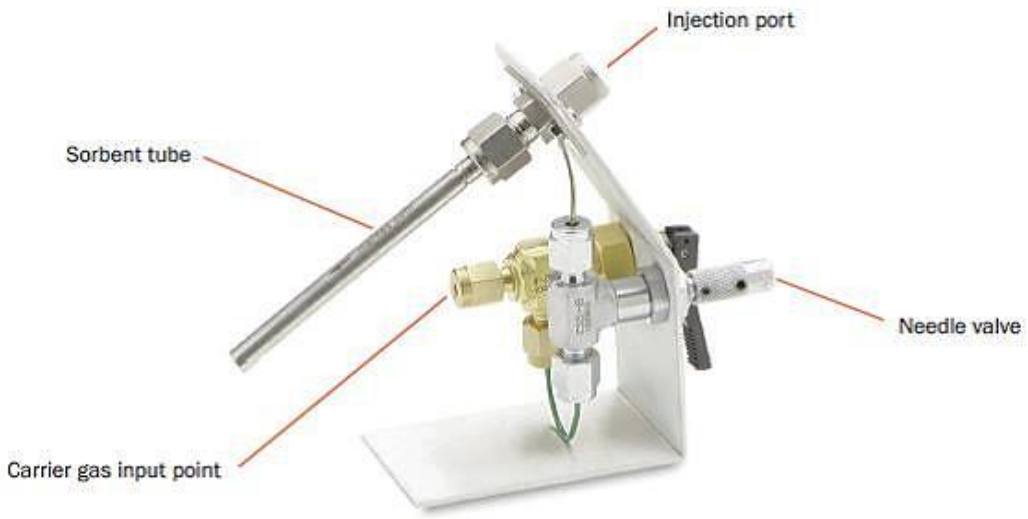


Figure 2. 8 – Calibration Solution Loading Rig (CSLR™), Markes International Ltd (MARKES, 2014).

The CSLR™ setup consists of an injection port with a controlled carrier gas supply and a sorbent tube connection. The sampling end (grooved end) of the sorbent tube is connected to the back side of the port. The carrier gas flow (typically 99.999% purity nitrogen or helium) is set with a needle valve to 50 – 100 mL min^{-1} (MARKES, 2014). The standard solution or gas standard is introduced through the injector septum using an appropriate precision syringe. Liquid standards of VOCs vaporise in the flow of gas, allowing the solvent (methanol) and analytes to reach the sorbent bed in the vapour phase.

2.6.2 Desorption of VOCs from TD tubes

The VOC mixture initially adsorbed onto the sorbent surface is then heated and back flushed with the carrier gas (helium); the compounds are released from the sorbent into the cold trap, which is usually maintained at -7 °C; the sample is transferred to the GC column by rapidly heating the cold trap and back flushed with helium; following trap desorption the trap is cooled down prior to the next desorption. Thermal desorption is usually carried out automatically by means of using a TD autosampler (Figure 2.9) connected to a thermal desorption unity.

Figure 2. 9 – ULTRA™ thermal desorption autosampler, MARKES International.

Split/splitless injections may be used, particularly for highly concentrated samples which the split injection is recommended. However, split injection is not useful for trace gas analysis.

2.6.3 Chromatographic separation

Chromatography is a physical separation process, where VOCs undergo separation along the GC column according to their relative affinity for the stationary phase of the column. The strength of the analyte-stationary phase; the vapour pressure, and the relative polarity of the analyte determine the order of elution of the VOCs. The lower the boiling point is, the higher the vapour pressure of the compound and the shorter retention time usually is, i.e. VOCs with higher vapour pressure, elute faster. If the polarity of the stationary phase and compound are similar, the retention time increases because the compounds are more strongly adsorbed and take longer to migrate through the column.

Chromatographic analysis with GC-MS generally uses fused silica capillary columns, liquid phase coated with the stationary phase, called Wall Coated Open Tube (WCOT) column (Sparkman *et al.*, 2011). The stationary phase of these columns is generally constituted of siloxane based polymers or polyethylene glycol (PEG).

The complex chemical diversity of VOCs in clinical samples requires the use of longer GC columns, typically a column length of 60 metres, improving separation of analytes along the column. These chemical characteristics make column selectivity, thermal stability, and inertness critical to resolving volatiles. Rxi®-624Sil MS columns offer reliable resolution of VOCs and also provide lower bleed and greater inertness than other columns. Column specifications, such as the column internal diameter, improves the separation of analytes and reduces the analysis time, but on the other hand reduces the column capacity. As regards to film thickness, increasing the film thickness improves separation of analytes, however it also increases the run time.

The GC column is kept in a heated oven. The separation of VOCs of different volatilities requires the use of temperature ramps, allowing the separation of compounds at different retention times. Whenever possible low bleed columns should be used, nevertheless, high oven temperatures operated over a long period may lead to the thermal decomposition of the column and consequently, column bleeding products will be detected and possibly will stain the ion source. Typically, the column bleed incorporates phenyl or phenyl-type group into siloxane polymer backbone structure, and *m/z* 207 is commonly known as column bleed product.

The carrier gas flow rate has a significant impact on the resolution. Apart from hydrogen, helium provides the best resolution for higher flow rates, improving the resolving power and shorter run times.

2. 7 Biomarkers vs. Biomarker Profiles

The complex relationships between a number of different compounds and the presence or absence of a disease or condition indicates that perhaps volatile biomarkers profiling with bioinformatics is a more promising approach. A specific breath marker related to a specific disease is the ideal. However, this is unlikely to be the case for the majority of diseases or conditions, where it is more probable that a range of VOCs with varying concentrations will have to be used. By adopting a strategy of identifying patterns, rather than trying to identify individual VOCs, “breath fingerprinting” could provide a suitable and reliable method for discriminating between healthy and diseased states. This approach requires elaborate methods of data analysis, pattern recognition techniques, such as principal component analysis (PCA) and partial least squares discriminant analysis (PLSDA). Principal component analysis is a mathematical algorithm that reduces the

dimensionality of the data. It accomplishes this reduction by identifying directions, called principal components, in which the variation in the data is maximal. Samples can be plotted, and visually assess similarities and differences between samples, and determined whether samples can be grouped. Other multivariate methods also exist, such as PLSDA, or support vector machines (SVMs). All of these methods use whole profiles, and yet it is possible to identify individual components (e.g., compounds or ions), which are most responsible for the differences observed between groups (e.g., groups of samples positive or negative for a disease). Thus, "biomarkers" (which are not usually unique) may be identified in this way. Cross validation of the models is used to predict the classification capabilities on unknown objects. Hence, if there is a good correlation between the predicted and actual values this means that the model fits.

In clinical practice, biomarkers such as genes and proteins are identified and quantified in order to track the biochemistry within the body for a specific disease. However, clear quantification of VOCs is a harder job due to the difficulties in finding the biochemical pathways in the body for each metabolite. This will involve a close collaboration between clinicians and analytical chemists. The need to understand the relationships between many variables makes multivariate analysis an inherently difficult subject. It is important to note that when the number of variables quickly overwhelms the number of samples, spurious correlations may be found (Miekisch *et al.*, 2012). Confounding variables may lead to wrong conclusions. Confounding variables comprise environmental compounds, physiological parameters and even the sampling procedures (Miekisch *et al.*, 2012). The statistical technique used to control the influence of confounding variables is called Analysis of covariance (ANCOVA). Classification of the subjects into groups is achieved by discriminant analysis, cluster analysis, and propensity score analysis. Clustering attempts to find similarities among the subjects that were measured instead of among the

measures that were made. For multiple dependent variables, in which two or more dependent variables are included, multivariate analysis of variance (MANOVA) and canonical correlation analysis are applied. Recently, Halbritter and co-workers (Halbritter *et al.*, 2012) used the MANOVA technique to discriminate according to whether the pregnant women had gestational diabetes mellitus, impaired glucose tolerance, or normal glucose tolerance, by means of analysing the women´s breath by PTR-MS and correlating it with the oral glucose tolerance test.

The success of proper statistical analysis is to have a good statistical validation, as well as trustworthy biological interpretation of the results.

Chapter 3

Materials and analytical/statistical methods

3. 1 A potential method for comparing instrumental analysis of breath gas volatile organic compounds using standards calibrated for the gas phase

Chapter 4 discusses a potential method for comparing instrumental analysis of breath gas volatile organic compounds found in breath using standards calibrated in the gas phase. Standards calibrated for the gas-phase were generated according to Henry´s law (see below), and these were analysed via Selected Ion Flow Tube Mass Spectrometry (SIFT-MS), Proton Transfer Reaction Mass Spectrometry (PTR-MS) and Gas Chromatography-Mass Spectrometry combined with thermal desorption (TD-GC-MS). Analysis by SIFT-MS and TD-GC-MS was carried out at The Open University (OU) and the solutions for PTR-MS headspace were transported from the OU to University of Birmingham in DURAN® glass bottles, so the same solution could be used for SIFT-MS and PTR-MS.

3.1.1 Samples

Accurate creation of partial pressures of volatile compounds in the headspace is an essential requirement for a correct determination of VOC concentration. Thus, this requires an understanding of liquid-phase/gas-phase equilibrium, commonly known as Henry's law. The Henry's coefficient, k_H, relates the concentration of the compound in a dilute aqueous solution and its headspace partial pressure so that at a constant temperature, the molar concentration of the compound in the liquid is directly proportional to its vapour pressure in the gas phase. Moreover, Henry's constant is temperature dependent, typically increasing with temperature at low temperatures (Smith and Harvey, 2007). To calculate Henry's law coefficients via this method the following equation (equation 3.1) is applied:

Equation 3.1

$$k_H = k_H^\circ \times exp\left(\frac{-\Delta_{vap}H}{R}\left(\frac{1}{T} - \frac{1}{T^\circ}\right)\right)$$

A correction can be made according to temperature (equation 3.2):

Equation 3.2

$$\frac{dlnk_H}{d(1/T)} = \frac{-\Delta_{vap}H}{R}$$

Here, k°_H represents the Henry's law constant for solubility in water at 298.15 K; $-\Delta_{vap} H/R$ is the temperature dependence parameter; T° is the standard temperature of 298.15 K; and T is the actual temperature. In this work, the values k°_H and $dlnk_H/d(1/T)$ are tabulated in Table 3.1.

Standard aqueous solutions of six VOCs acetone, ethanol, methanol, 1-propanol, 2-propanol and acetaldehyde (which are physiologically important and present in human breath), were created to produce headspaces containing known concentrations of these compounds in the vapour phase, as calculated using the measured temperature and Henry's law coefficient (k_H) (Table 3.1). Mean values were calculated for $k°_H$ and $\Delta_{vap}H/R$ and these values were empirically determined (Sander, 1999).

Table 3. 1 – Henry's law constants at 298 K (k°H), ΔvapH/ R values in K and the derived Henry's law constants at 293 K for acetone, ethanol, methanol, 1-propanol, 2-propanol, acetaldehyde in aqueous solution. Mean values are given for k°H and ΔvapH/ R.

	$k°_H$ [(mol dm^{-3}) atm^{-1}] 298 K [a]	$\Delta_{vap}H/R$ [K] [a]	k_H [(mol dm^{-3}) atm^{-1}] 293 K
Acetone	30.0	4600	39.0
Ethanol	184.4	6500	267.5
Methanol	203.8	5400	277.5
1-propanol	138.0	7500	211.9
2-propanol	127.0	7500	195.0
Acetaldehyde	12.9	5371	17.5

[a] Reference (Sander, 1999)

Henry's law coefficients (k_H) at 293 K were calculated for each VOC according to equations 3.1 and 3.2. In general, the higher the partial pressure of a species in the gas-phase under equilibrium conditions, the lower the Henry's law coefficient (k_H), since the aqueous-phase/gas-phase equilibrium is defined as $K_H = C_a/p_g$, where C_a is

aqueous-phase concentration of a particular species, and p_g is the partial pressure of that species in the gas phase (Sander, 1999).

Aqueous solutions were prepared using accurate micropipettes and calibrated for the headspace (10 ppm) at 293 K. Individual one litre solutions (10 ppm) were prepared as follows: 29 µL acetone, 156 µL ethanol, 112 µL methanol, 158 µL 1-propanol and 149 µL 2-propanol were added to individual clean glass bottles respectively, and the final volume (one litre) adjusted with purified (deionised) water. These solutions were used to provide more dilute solutions. Diluted solutions were prepared individually, the volumes 250 mL, 50 mL and 5 mL were added to 500 mL glass bottles to provide more dilute solutions that were expected to give headspace concentrations of 5 ppm, 1 ppm and 0.1 ppm respectively. The 500 mL volume was adjusted using purified (deionised) water. A more concentrated solution of acetaldehyde (1000 ppm) was prepared, where 1000 µL of acetaldehyde were added to a clean glass bottle containing 1 litre of purified (deionised) water. This concentrated solution (1000 ppm) was used to provide more dilute solutions of acetaldehyde. Diluted solutions were prepared individually, the volumes 5000 µL, 2500 µL, 500 µL and 50 µL were added to 500 mL glass bottles to provide more dilute solutions that were expected to give headspace concentrations of 10 ppm, 5 ppm, 1 ppm and 0.1 ppm respectively.

The reproducibility of standard generation was determined by preparing five solutions of 2-propanol 10 ppm in the headspace, and these solutions were analysed via SIFT-MS. Thus, the standards were prepared with a coefficient of variation determined at 6%. The coefficient of variation is defined as the ratio of the standard deviation to the mean.

For the analysis by SIFT-MS, the calibration standards were placed (5 mL of each), individually, inside a 40 cm long Nalophan sampling bag, made up of 65 mm diameter Nalophan NA tubing 25 μm thick (Kalle UK). Regarding the analysis by PTR-MS, the same procedure was used, although, in this case the sample was placed inside a 45 cm long Nalophan sampling bag, made up of 135 mm diameter Nalophan NA tubing, 25 μm thick (Kalle UK). All the sample bags were sealed and filled with hydrocarbon free air (Air Products) to generate the VOC headspace, or N_2 (BOC Gases, UK) for the PTR-MS measurements respectively.

The headspaces were allowed to equilibrate at ambient temperature (where the temperature was carefully measured and recorded for applying enthalpy correction) for 15 minutes after which the sampling bags were tested via SIFT-MS, in the case of PTR-MS sampling took place after 60 minutes. These differences are due to the equilibration time depending on the size of the bag, thus there is the need to guarantee that VOC are in equilibrium in the gas-phase, and will produce the accurate headspace concentration no matter the size of the bag. Therefore every headspace profile were tracked using the multiple ion monitoring (MIM) mode capability of SIFT-MS to ensure equilibrium in the gas phase. In the PTR-MS instrument a sample bag was tracked over time and its concentration reached equilibrium at 60 minutes. The reproducibility of the sampling procedure was evaluated by preparing five sampling bags (40 cm long Nalophan sampling bag, made up of 65 mm diameter Nalophan NA tubing 25 μm thick) containing 5 mL of the same solution (2-propanol 10 ppm in the headspace) calibrated for the headspace. Thus, the reproducibility of the sampling procedure was determined by determining the coefficient of variation which in this case was 3%.

For GC-MS analysis the VOCs were pumped for 5 min at a constant flow of 100 mL min^{-1} into pre-conditioned stainless-steel TD sorbent tubes, according to the detailed description given in section 3.1.4. VOC mixtures were prepared according to Henry´s law in the same manner. The chemicals were mixed to give solutions of approximately (although accurately calculated) 0.1 ppm, 1 ppm, 5 ppm and 10 ppm. Then 5 mL of each was placed in 40 cm long Nalophan sampling bags. The bags were sealed and filled with hydrocarbon free air (Air Products) and the generated headspaces were left to equilibrate for 15 minutes. Aqueous liquid headspace sampling was performed using SIFT-MS, PTR-TOF-MS and TD-GC-MS.

All the chemicals used in this study were purchased from Sigma-Aldrich (purities ≥ 99%) and all were used without any further purification.

3.1.2 SIFT-MS

Data were collected using the Mk2 instrument (PDZ Europa, UK) with a flow rate corresponding to a pressure of 0.008 Torr. The sample bags were attached to a heated capillary and further analysed. Full spectra of the count rates at each *m/z* value in the range *m/z* 10 to *m/z* 140 were recorded for all the samples via the precursor ion H_3O^+, using the full scan mode and a total sample time of 30 seconds. The liquid-phase/gas-phase equilibrium was confirmed for each sample by means of using the multiple ion mode (MIM). Concentrations of the volatiles were automatically determined under *thermal conditions* using an on-line kinetics database containing reaction rate coefficients, developed from numerous detailed selected ion flow tube studies of various classes of compounds (alcohols, aldehydes, ketones, hydrocarbons, *etc.*) with the three precursor ions (Smith and Španěl, 1996b). Background laboratory air was also analysed.

3.1.3 PTR-MS

Measurements were performed with a commercial PTR-TOF-MS (Kore Technology Limited) equipped with an ion funnel, although for this study the ion funnel facility was not used. The drift tube length is 9.04 cm. This instrument is particularly suited for reactions between the precursor ion H_3O^+ and gas samples. The drift tube working conditions (pressure and reduced electric field) were set to deliver the highest signal of protonated product ion while the drift tube temperature (373.15 K) was kept constant over the experiments. The drift chamber parameter settings (voltage, pressure and *E/N*, where *E* is the electric field strength and *N* is gas number density in the reaction chamber; Td = Townsend, where 1 Td = 10^{-17} cm^2 V) are listed in Table 3.2.

Table 3. 2 – Ionization conditions in the drift tube of PTR-TOF-MS.

VOC	***Drift voltage*** *(V)*	***Drift pressure*** *(mbar)*	***E/N*** *(Td)*
Acetone	366	1.30	125
Ethanol	230	1.31	99
Methanol	380	1.37	125
1-propanol	285	1.38	94
2-propanol	209	1.31	90
Acetaldehyde	370	1.31	160

Mass spectra were typically collected in the range *m/z* 20 to *m/z* 250 with an integration time of 1 minute. Background laboratory air was also analysed. The concentrations of the volatiles were determined through the theoretical prediction of reaction rate coefficients.

3.1.4 TD-GC-MS

After headspace analysis by SIFT-MS, VOCs were pumped into pre-conditioned stainless-steel TD sorbent tubes for 5 minutes at a constant flow of 100 mL min^{-1} (Pump TSI Inc. SidePak Model SP730). These tubes had dual packing comprising of 50% Tenax and 50%

Carbotrap (Markes International Limited). Prior to analysis, the tubes were spiked with 1 µL of internal standard, d8-toluene in methanol, and were then flushed with nitrogen for 30 seconds. The same protocol was used for the analysis of breath- and skin VOCs trapped in TD tubes. All samples were analysed in random order according to the method described below.

Chromatographic analyses were performed using an Agilent 6890/5973 GC-MS system (Figure 3.1) equipped with a Markes TD autosampler, and Markes UNITY thermal desorber. The tubes were submitted to a pre-purge of 1.0 minute, followed by desorption at 260 °C for 3.0 minutes. The trap temperature was set at -7 °C and the actual trap desorption occurred at 300 °C for 3.0 minutes. The volatiles of interest were separated using a Restek column (60 m × 0.32 mm, film thickness 1.8 µm) working in a constant flow mode (helium at 30.8 mL min^{-1}). The column temperature program involved an initial increase from 35 °C to 60 °C at a rate of 11 °C min^{-1}, followed by a rate of 20 °C min^{-1} up to 220 °C, and a constant temperature of 220 °C for 10 minutes. The mass spectrometer was operated in a SCAN mode with an associated *m/z* range set from 33.0 to 260.0. The quadrupole, ion source, and transfer line temperature were kept at 150 °C, 230 °C and 230 °C, respectively.

Figure 3. 1 – Mk2 instrument SIFT-MS (PDZ Europa, UK) and Agilent 6890/5973 GC-MS system equipped with a Markes TD autosampler, and Markes UNITY thermal desorber.

3.1.5 Data processing

Prior to data analysis, the counts per second acquired by SIFT-MS were normalised against the counts per second of the H_3O^+ (*m/z* 19) precursor ion. A high counts per second ($\leq 10^7$ cps) for H_3O^+ at *m/z* 19 occurs using PTR-MS, therefore detector saturation follows. For this reason, the isotope of oxygen, ^{18}O, is commonly used i.e. the, count rates of $H_3{}^{18}O^+$ at *m/z* 21 were measured to assess the counts per second of H_3O^+. The natural abundance of ^{18}O relative to ^{16}O is approximately 1:500, thus count rates at *m/z* 21 were multiplied by 500 to obtain the actual count rates for H_3O^+. Cluster ions [$H_3O^+(H_2O)_n$, n = 1, 2] at *m/z* 37 and *m/z* 55 were taken into account in the determination of H_3O^+ count rates. PTR-MS counts per second have been normalised to a H_3O^+ signal of one million counts per second, i.e. $ncps = 10^6 \times i(MH^+)/i(H_3O^+)$, where MH^+ is the protonated parent ion.

Using the internal standard approach via GC-MS, the determination of VOC concentration expressed in ng L^{-1} was achieved by means of the following equation (equation 3.3):

Equation 3.3

$$C = \frac{abundance\ (\%)_{VOC}}{abundance\ (\%)_{d8-toluene}} \times 50\ ng\ \times \frac{1000\ ml}{100\ ml\ min^{-1} \times 5\ min\ \times 1\ L}$$

where, the VOC concentration (*C*) is expressed in ng L^{-1} and the concentration expressed in ppbv was calculated; relative abundances of VOC and d8-toluene are expressed in percentage (%); 50 ng represents the mass of d8-toluene used to prepare the standard; with representative VOCs being pumped into sorbent tubes for 5 minutes at a constant flow of 100 mL min^{-1}.

GC-MS data analysis was performed through the aid of *AMDIS* (*Automated Mass Spectral Deconvolution and Identification System*) software, and followed by reliable identification of compounds using the *NIST* (*National Institute of Standards and Technology*) library.

3.1.6 Product ions and branching ratios

All of the analytes in this study have higher proton affinities than H_2O (691.0 kJ mol^{-1}) (Linstrom and Mallard, December 2013), and so rapid proton transfer occurs in each case. The low molecular weight alcohols have a proton affinity range of 754.3–793.0 kJ mol^{-1} (Linstrom and Mallard, December 2013). Acetaldehyde proton affinity is 768.5 kJ mol^{-1} (Linstrom and Mallard, December 2013), while for acetone it is 812.0 kJ mol^{-1} (Linstrom and Mallard, December 2013). In this study proton transfer from H_3O^+ to VOCs has been investigated with the product ions and their associated branching ratios being identified in chapter 4 (Table 4.1). The standard aqueous solutions prepared are highly humid as would be expected if aqueous solutions are present with a small enclosed headspace, such as a sample bag. Therefore, cluster ions [e.g., $H_3O^+(H_2O)_n$, n = 1, 2, 3] can form from the precursor ions and these cluster ions can also act as precursor ions and lead to hydrated product ions $RH^+(H_2O)_n$, n = 1, 2, 3 (Španěl and Smith, 2000a). These hydrated product ions must be taken into account when determining VOC concentrations. Accordingly, and as expected, the SIFT-MS instrument yielded much higher abundances of cluster ions compared to PTR-MS since the average kinetic energy of the ions are elevated above thermal in PTR-MS over the entire length of the drift tube, consequently, cluster ion formation is effectively prevented. Due to the absence of an electric field in SIFT-MS it is therefore possible to carry out ion-molecule reactions under thermal conditions, where the kinetic behaviour is well known, and less fragmentation of product ions occurs. The fragmentation patterns observed at specific *E/N* are further described below in chapter 4.

Our findings indicated that quantification via PTR-MS is sensitive to alterations in the reduced electric field (i.e. changes in the *E/N* ratio), therefore, the drift tube working conditions (pressure and reduced electric field) were set to deliver the highest signal of protonated product ion, and this was found to deliver the best approximation to the actual gas-phase concentration inside the Nalophan bag.

Product ions and branching ratios found in literature for the VOCs in study now follows.

Methanol: This common breath metabolite methanol has been thoroughly studied using SIFT-MS and PTR-MS (Španěl and Smith, 1997, Brown *et al.*, 2010). Our results are in agreement with earlier measurements, where non-dissociative proton transfer predominantly occurs either for SIFT-MS and PTR-MS. The reaction by SIFT-MS yielded protonated methanol ROH_2^+ (41%) at *m/z* 33 and the water-cluster ions $ROH_2^+(H_2O)_{1,2}$. The reaction at an *E/N* of 125 Td for PTR-MS mainly yielded the protonated ion at *m/z* 33 and clustering of the protonated methanol $ROH_2^+ \cdot H_2O$ at *m/z* 51.

Acetaldehyde: Saturated aldehyde reactions with H_3O^+ have been previously studied by SIFT-MS, where only non-dissociative proton transfer is seen (*m/z* 45) for acetaldehyde reactions (Španěl *et al.*, 1997, Mochalski *et al.*, 2014b). Cluster ions $ROH^+(H_2O)_{1,2}$ are also observed. At 300 K there is an increasing tendency to eject water molecules for C_4 and higher species (Španěl *et al.*, 1997). Recently, product ion distributions for the reactions of NO^+ with twenty two aldehydes under dry and humid conditions were carried out using a selective reagent ionisation time-of-flight mass spectrometer (SRI-TOF-MS) at a specific *E/N* of 130 Td (Mochalski *et al.*, 2014b). However ion chemistry studies on H_3O^+–acetaldehyde reactions by PTR-MS have not yet been exhaustively studied.

In this work, analysis of acetaldehyde was carried out at *E/N* of 160 Td, where non-dissociative proton transfer is seen, yielding *m/z* 45 as the major product ion, with additional clustering of the protonated ion to produce $ROH^+ \cdot H_2O$ at *m/z* 63. In another research group was found that H_3O^+-acetaldehyde studied at *E/N* of 112 Td, produces only the protonated ROH^+ plus H_2O (Jordan *et al.*, 2009a).

Ethanol: Similarly to methanol, the reaction of ethanol through SIFT-MS yielded the protonated ion ROH_2^+ at *m/z* 47 and was observed an additional clustering of the protonated ethanol monomer species to H_2O to produce $ROH_2^+(H_2O)_{1,2}$, which is in agreement with earlier studies developed by Španěl *et al.* (Španěl and Smith, 1997). For our measurements with PTR-MS, an *E/N* of 99 Td was used and the reaction yielded ROH_2^+ at *m/z* 47 as the major product ion, and additional clustering of the protonated ion to produce $ROH_2^+ \cdot H_2O$ at *m/z* 65. Our findings are in agreement with the measurements of Brown *et al.* for a low *E/N*, where we observed traces of the fragment ion at *m/z* 45 ($C_2H_5O^+$) corresponding to a loss of H_2 (Brown *et al.*, 2010).

Acetone reacts by non-dissociative proton transfer to produce a single protonated product ion ROH^+ at *m/z* 59 either by SIFT-MS and PTR-MS. An additional water-cluster ion $ROH^+ \cdot H_2O$ at *m/z* 77 is yielded in the flow tube of SIFT-MS (Španěl *et al.*, 1997).

1-Propanol and 2-Propanol: The branching ratios for such isomers have been thoroughly studied by Warneke *et al.* who determined that the proton transfer between H_3O^+ and a VOC is non-dissociative (Warneke *et al.*, 1996).

The isomeric compounds 1-propanol and 2-propanol show a similar behaviour via SIFT-MS and PTR-MS. In earlier findings in PTR-MS, it was shown the loss of water from the

protonated alcohol is predominant and, consequently, breaking up to *m/z* 43 ($C_3H_7^+$) (Brown *et al.*, 2010, Španěl and Smith, 1997). These isomers showed a similar fragmentation pattern for a low *E/N* (Brown *et al.*, 2010). Both yield a major fragment ion at *m/z* 43 ($C_3H_7^+$), except for the fragment ion at *m/z* 59 ($C_3H_7O^+$) which in particular is observed for 2-propanol but not from 1-propanol. The fragment ion at *m/z* 59 originates through the loss of H_2 from the protonated 2-propanol (Brown *et al.*, 2010).

3. 2 Analysis of VOC profile in faecal headspace, and exhaled breath and skin from type-2 diabetic mice

This section concerns the study of volatile organic compounds (VOCs) coming from different mice models for type 2 diabetes. Two different studies were carried out, a longitudinal study of the VOC profile emitted by the faecal headspace of Cushing´s syndrome mouse model of type 2 diabetes, and a longitudinal study of the VOC profile emitted by the faecal headspace of single *Afmid* knockout mice exhibiting impaired glucose tolerance. Faecal headspace was analysed by SIFT-MS and TD-GC-MS and related to changes in blood glucose in type 2 diabetes (T2D) and impaired glucose tolerance respectively. A cohort of 65 Cushing´s mice (21 male and 44 female) were bred, and a cohort of 48 *Afmid* animals (24 male and 24 female) were bred.

Faecal samples and breath samples were collected by myself at the Mary Lyon Centre (Harwell, UK) as well as blood glucose analysis was performed by myself at the Mary Lyon Centre. ELISA (Enzyme-Linked Immunosorbent Assay) measurements were performed either by Dr Liz Bentley or Dr Marianne Yon at the Mary Lyon Centre. Faecal samples were then transported to The Open University (OU) in dry ice and kept at -80 °C until further analysis. Data were acquired using SIFT-MS and TD-GC-MS at OU through the use of the protocols described in section 3.1.4 and 3.2.6.

All animal studies were carried out using guidelines issued by the United Kingdom (UK) Medical Research Council (MRC), in Responsibility in use of Animals for Medical Research (July 1993) and Home Office Project License number 30/3146. All animal studies were carried out at the Mary Lyon Centre (Harwell, UK).

Mice (Cushing's mice, *Afmid* mice and wild-type littermates) were maintained in controlled light (12 h light and dark cycle), temperature (21 ± 2 ˚C) and humidity (55 ± 10%). Mice had free access to water (10–12 ppm chlorine) and were fed *ad libitum* on a commercial diet (5.3% fat [corn oil], 21.2% protein, 57.4% carbohydrate, 4.6% fibre; Rat and Mouse Diet No. 3 (RM3), Special Diet Services, Essex, UK). All mice were housed in groups of 2–5 mice.

Data processing: Prior to statistical analysis, the intensity of *m/z* values (SIFT-MS) within the data were normalised against the intensity values of the H_3O^+ (*m/z* 19) precursor ions. All *m/z* values were included in the statistical analysis for H_3O^+ precursor ion. The known adduct ions, *m/z* values: 19, 21, 30, 32, 34, 37, 39, 48, 55, 57, 66, 73, 75, and 91, were included for all data sets (H_3O^+, NO^+ and O_2^+). In this study assessing the faecal headspace levels, 18 animals (Cushing´s mice) have been removed from the SIFT-MS NO^+ data set obtained at 12 weeks due to a read fault on the SIFT-MS software. Whereas, 18 animals (Cushing´s mice) have been removed from the SIFT-MS O_2^+ data set obtained at 12 weeks due to the same issue. Only one animal was removed from the SIFT-MS O_2^+ data set for Cushing´s mice at 16 weeks of age.

GC-MS data analysis was performed according to the detailed description given in 3.1.5.

The statistical analysis presented here was performed using the software *IBM SPSS Statistics 21.0*.

Three mice models were used in this study: Cushing's mice, *Afmid* mice and wild-type littermates. Details of these are given below.

3.2.1 Mouse model of Cushing´s Syndrome

These mice were produced through the use of ENU mutagenesis described in chapter 1.

Breeding strategy: Male Heterozygous Crh $^{-120/+}$ mice on a B6-C3PDE F1 background were backcrossed to B6-C3PDE F1 female wild-type mice to produce progeny which were screened for phenotypes and classified as mutant (het) mice and wild-type (WT) mice. A cohort of 65 animals (21 male and 44 female) was used of whom 30 were genotyped as mutants and 35 WT littermates (Aigner *et al.*, 2008). Mapping and sequencing (Aigner *et al.*, 2008) was performed at Mary Lyon Centre.

3.2.2 Single Afmid knockout mice

These mice were produced through gene knockout described in chapter 1 (Liu *et al.*, 2003).

Breeding strategy: Wild-type (*Afmid*$^{+/+}$) and knockout (*Afmid*$^{tm1b/tm1b}$) mice on a C57BL6/NTac background were generated by inter-crossing *Afmid*$^{tm1b/+}$ mice. A cohort of 48 animals (24 male and 24 female) were bred of whom 30 were genotyped as mutants (hom) and 18 WT littermates. It has been found that mice homozygous (hom) for the tm1b deleted allele have impaired glucose tolerance (Hugill *et al.*, 2015).

3.2.3 Metabolic cages

Stool samples were collected from each mouse at various times for analysis of volatile organic compounds (VOCs). Animals were individually housed in metabolic cages

(Techniplast Kettering, UK) (Figure 3.2) for 30 minutes or until sufficient gathering of stool, i.e. six pellets per animal. The mice were fed *ad libitum* on water (10–12 ppm chlorine) and powdered chow (RM3, Special Diet Services Essex, UK). Autoclavable red transparent igloos (Datesand Ltd, Manchester, UK) were provided as environmental enrichment. Faecal samples were collected into Eppendorf containers, and these were immediately frozen at -20 °C.

Cushing´s mice were tracked over 4 months, thus, samples were collected at 8, 12, 16 and 20 weeks of age respectively. *Afmid* mice were tracked over 3 months, therefore, samples were collected at 8, 12 and 16 weeks of age respectively.

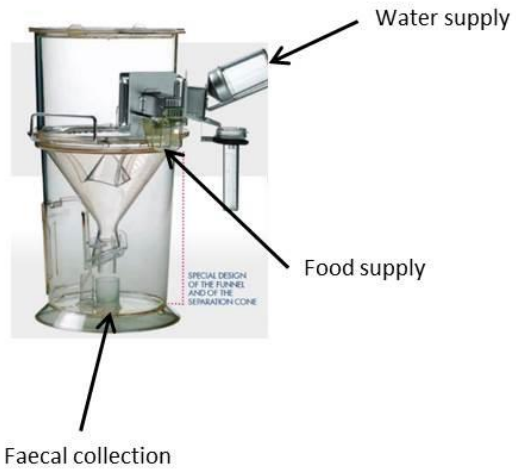


Figure 3. 2 – Metabolic cage Techniplast Kettering, UK (Adapted from TECNIPLAST).

3.2.4 Blood and plasma biochemistry

Animal weight was recorded prior to analysis and over the duration of the study. Blood samples were taken on the same days as faecal samples and were taken from the lateral tail vein following application of topical local anaesthesia (EMLA cream) leaving it to work for 15 minutes, and while mice were restrained in a Perspex bleeding tube. Blood glucose levels were measured at the same period of the day, specifically between 2 pm and 4 pm, using an AlphaTRAK2 blood glucose meter. Cushing´s mice blood glucose was measured

over 4 months, i.e. at 8, 12, 16 and 20 weeks old. *Afmid* mice blood glucose was measured over 3 months, i.e. at 8, 12 and 16 weeks of age.

In order to evaluate the insulin levels in mice, plasma concentrations in blood were measured after terminal bleeding of *Afmid* and Cushing´s mice. An ELISA kit (Cat no. EZRMI-13K) was used and the range of the assay set between 0 and 15 ng mL^{-1}. ELISA measurements were performed either by Dr Liz Bentley or Dr Marianne Yon at the Mary Lyon Centre (Harwell, UK). Detailed information may be found in the Appendices.

3.2.5 Offline breath- and skin analysis in unrestrained mice

Offline breath analysis from unrestrained Cushing´s mice was also attempted and VOCs coming from breath/skin were trapped in TD tubes and further analysed by SIFT-MS and TD-GC-MS. However, the experiments did not work out as we expected and no results are presented in this thesis. This work has, therefore confirmed that offline breath analysis in unrestrained mice is challenging, and I faced many difficulties during the set-up of the methodology, which will be listed here. Direct sampling was not possible, thus the breath VOC sampling was performed directly into TD tubes, and detection of VOC at trace concentration with enough sensitivity was problematic. In addition, it is difficult to have control over the contamination of the VOC sample, i.e. mice tend to urinate and let go stools under unfamiliar or stressful situations, not to mention an increase in physical activity. Nevertheless, the suggested methodology is indicated below.

A plethysmography chamber, commonly used for measuring the respiratory function in unrestrained mice, was used in an attempt to collect VOCs being emitted from breath- and skin in unrestrained Cushing´s mice. The chamber was connected to a hydrocarbon free air cylinder (Air Products) and the flow set to 60 mL min^{-1}. The chamber was

prepared with a small orifice acting as pressure discharge. The empty chamber was flushed with clean air prior to the measurements for 3 minutes; a blank sample was taken; the mouse was placed in the chamber and this was flushed once again with clean air for 2 minutes; Air samples were then pumped (Gilian, GilAir Plus) into stainless-steel thermal desorption (TD) tubes [50% Tenax and 50% Carbotrap (Markes International Limited)] at 50 mL min^{-1} for 5 minutes. Duplicates were taken for analysis by SIFT-MS and TD-GC-MS. TD sorbent tubes were transported to OU and kept in the freezer at -20 °C for 1 week until further analysis.

TD tubes were analysed via SIFT-MS by means of using a prototype thermal desorber developed by Kore Technology Limited (Figure 3.3).

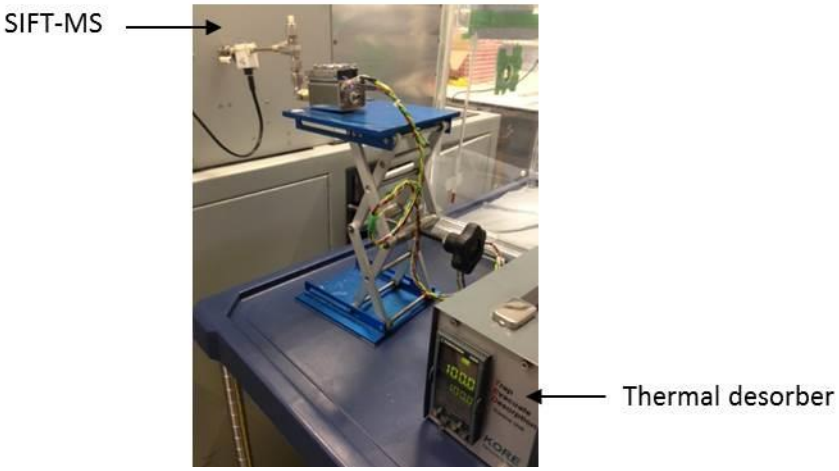


Figure 3. 3 – Direct desorption into SIFT-MS using a prototype thermal desorber developed by Kore Technology Limited.

The device (Figure 3.3) consists of a ceramic in which the tube is placed; the tube is gradually heated at different temperature ramps ranging from 20 °C to 200 °C and controlled by a simple temperature controller. The VOC are therefore desorbed from the stainless-steel tube into the SIFT-MS directly. The SIFT-MS was operated in MIM mode acquiring data for acetone, propanol, ethanol, pentanol, acetic acid and propanoic acid.

3.2.6 Faecal headspace analysis

Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) technique was discussed in detail in chapter 2 as a technique suited for the analysis of VOCs in real-time with high sensitivity and specificity. The methodology used for the analysis of faecal headspace (Figure 3.4) by SIFT-MS is discussed below. After the faecal headspace analysis by SIFT-MS, VOCs were pumped into pre-conditioned stainless-steel TD sorbent tubes and analysed by TD-GC-MS according to the description given in 3.1.4.

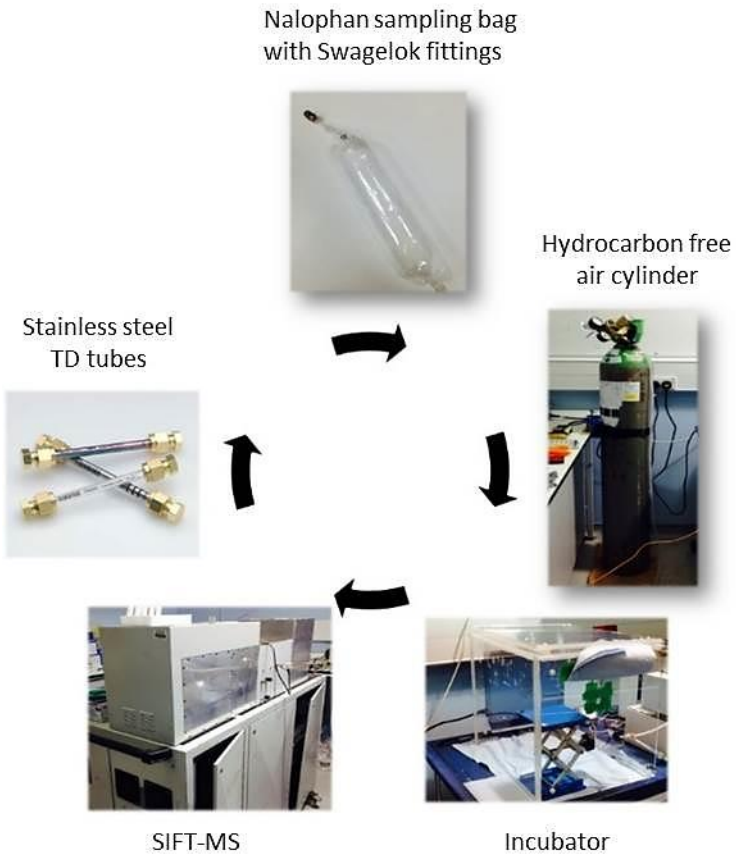


Figure 3. 4 – Illustrative diagram showing the methodology used for the analysis of VOCs via SIFT-MS and TD-GC-MS.

Six pellets per animal were placed inside Nalophan bags (35 cm long Nalophan sampling bag, made up of 65 mm diameter Nalophan NA tubing 25 µm thick (Kalle UK)); sample bags were sealed and filled with hydrocarbon free air (Air Products); after an 1 hour of incubation at 37 °C, the bags were connected to the SIFT-MS via the heated sampling line and the faecal headspace further analysed. The normalisation of the data corrects for the

variation in sample weight. Samples were analysed in random order and full spectra of the count rates in the range *m/z* 10 to *m/z* 140 were recorded for all the samples via H_3O^+, NO^+ and O_2^+, using the FS mode and a total sample time of 30 seconds. Background lab air was also analysed.

3. 3 Analysis of the volatile faecal metabolome in screening for colorectal cancer

3.3.1 Sample acquisition

The cancer samples in this study were collected through the National NHS Bowel Cancer Screening Scheme. All the subjects received a FOBT kit at home. A sample was taken on three different days. The subjects who tested positive through FOBT screening subsequently underwent a colonoscopy and were invited to participate in this study. These patients were requested to produce a further stool sample to bring with them in a wide plastic sterile container provided by the hospital. The samples were kept frozen at -80 °C and these were then shipped to The Open University on ice.

After the colonoscopy, the patients were classified into risk groups and the results are shown in Table 3.3.

Table 3. 3 – Risk class applied to histological results.

Class	***Detail***	***Number of samples***
1	Normal, nothing of concern found = **low risk**	31
5	High Grade Adenoma = **high risk**	31
6	Adenocarcinoma = **high risk**	

Initially, all the subjects tested positive for FOBT but this evidence was rapidly disproven by colonoscopy results. In this study, 31 samples were analysed from the 'normal' or low risk group (class 1) and 31 samples from the high risk groups (class 5 and 6), which

actually had colorectal cancer. The samples were stored at -80 °C until the moment of analysis by SIFT-MS.

3.3.2 Ethics statement

Samples were taken under the National NHS bowel cancer screening scheme. Favourable ethical opinion was obtained for subsequent inclusion of suitable subjects in this research by Cambridgeshire 2 REC, REF 08/H0308/13, and written informed consent obtained from each subject and consent form then kept securely. All samples were anonymised prior to analysis and only the cancer status was known to the authors (Batty *et al.*, 2015)

3.3.3 VOC sampling and analysis

Sample preparation: Faecal samples were removed from the -80 °C freezer and defrosted for 1 hour; 5 g of sample was directly introduced into a Nalophan bag (40 cm long Nalophan sampling bag, made up of 65 mm diameter Nalophan NA tubing 25 µm thick (Kalle UK)); the sample bags were sealed and filled with hydrocarbon free air (Air Products); the samples were incubated for 45 minutes at 37 °C.

SIFT-MS analysis: Data were acquired according to the detailed description given in 3.1.2. (PDZ Europa, UK) Full spectra of the count rates at each *m/z* value in the range *m/z* 10 to *m/z* 140 were recorded for all the samples via H_3O^+, NO^+ and O_2^+ using the full scan mode and a total sample time of 30 seconds. Background lab air was also analysed.

Data processing: Prior to statistical analysis, the intensity of *m/z* values within the data were normalised against the intensity values of the H_3O^+ (*m/z* 19) precursor ions. All *m/z*

values were included in the statistical analysis including the known adduct ions, *m/z* values: 19, 21, 30, 32, 34, 37, 39, 48, 55, 57, 66, 73, 75, and 91.

The statistical analysis presented in this thesis was performed using the software *IBM SPSS Statistics 21.0* and will now be discussed.

3. 4 Univariate Analysis

In order to perform a proper statistical analysis, the identification of sources of bias should be carefully taken into account and identified. Therefore, identification of outliers, i.e. scores very different from the rest of the data, and identification of the distribution of scores along the data set, i.e. spotting normality.

3.4.1 Kolmogorov-Smirnov statistic and Shapiro-Wilk statistic test

The Kolmogorov-Smirnov test and Shapiro-Wilk test evaluate the normality of the distribution of scores along the data set (Field, 2013). These tests compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation. If the test is significant ($p < 0.05$), then the distribution is significantly different from a normal distribution, thus indicating that the data set is not normally distributed. Inspection of normality is supported by the analysis of the *Normal Q-Q Plot*, where each score is plotted against an expected value from the normal distribution. A straight line suggests a normal distribution.

Considering a data set not normally distributed, non-parametric statistical techniques would have to be used from here on, or alternatively, transforming data not normally distributed into normally distributed might be one way of dealing with the matter.

3.4.2 T-test

The use of a T-test (independent samples) assumes a normally distributed set of scores, and this test looks for the differences between two groups, i.e. compares the mean scores between two different groups of people (e.g. males and females) or condition (e.g. disease and healthy) on the same continuous, dependent variable (Field, 2013). A significant result ($p \leq 0.05$) indicates that there is a significant difference between the groups on that dependent variable.

3.4.3 One-way repeated measures ANOVA with Bonferroni post-hoc test

In one-way repeated measures ANOVA, each patient is exposed to two or more different conditions, or measured on the same continuous scale on three or more occasions. This statistical technique assumes a normal distributed population and tells if there is a significant difference somewhere among the three set of scores (Pallant, 2013). A significant result ($p \leq 0.05$) indicates that there is a significant difference somewhere among the set of scores. Bonferroni *post-hoc* test is a correction to the statistical testing used to minimize the number of false discoveries that occur when conducting multiple tests. A Bonferroni adjusted p-value is used (p = 0.05/number of tests) accordingly (Pallant, 2013)

3.4.4 Mann-Whitney U test

The Mann-Whitney U test is a non-parametric technique and is used to compare differences between two independent groups, e.g. disease and healthy (Field, 2013). The non-parametric techniques assume a not normally distributed population. The Mann-Whitney U test is often considered the non-parametric alternative to the independent T-test. The Mann-Whitney U test uses the rank of each case to test whether

the ranks for the two groups differ significantly. A significant result ($p \leq 0.05$) indicates that there is a significant difference between the groups.

3.4.5 Spearman´s rho correlation coefficient

The Spearman´s rho correlation coefficient is a non-parametric statistic based on ranked data, so it can help to minimize the effects of extreme scores. Spearman's correlation determines the strength and direction (positive or negative) of the relationship between the two variables. Small correlations are indicated by rho´s correlation coefficient between 0.10 to 0.29; medium correlations are indicated by a correlation coefficient between 0.30 to 0.49; and large correlations present a correlation coefficient between 0.50 to 1.0 (Pallant, 2013). The rho´s correlation coefficient is strongly influenced by the size of the sample. Thus, in a small sample (e.g. n = 30) moderate correlations may not reach statistical significance at the traditional $p < 0.05$ level (Pallant, 2013). Spearman´s rho correlation was used in chapter 5 to determine the relationship between counts of *m/z* values and blood glucose levels in mouse models of diabetes.

3.4.6 Friedman test (repeated measures)

The Friedman test is the non-parametric alternative to the one-way ANOVA with repeated measures. This test is used when the same sample of participants are measured at three or more points in time, or under three different conditions. Statistically significant differences in the mean ranks for the three set of scores are indicated by p = 0.000 (which really means less than 0.0005), meaning that there is a statistically significant difference somewhere among the three set of scores (Pallant, 2013).

3.4.7 Wilcoxon Signed Rank test using a Bonferroni post-hoc test

Having established that there is a statistically significant difference somewhere among the three time points after using Friedman test, individual Wilcoxon Signed Rank tests combining a Bonferroni *post-hoc test* were used to determine which groups differ from the rest. A Bonferroni adjusted p-value is used (p = 0.05/number of Wilcoxon tests) accordingly (Pallant, 2013). Wilcoxon Signed Rank tests is a non-parametric technique and is used when the participants are measured on two occasions, or under two different conditions, thus it compares ranks at different time points. The effect size for this test can be calculated using the z value by the square root of N (Pallant, 2013).

3. 5 Multivariate statistics

Multivariate statistical analysis looks for relationships among multiple variables at the same time, and it looks for patterns (relative differences) rather than individual variables.

3.5.1 Discriminant analysis

Discriminant analysis was carried out between the groups (i.e. het/WT and hom/WT in chapter 5; Low Risk group/High Risk group in chapter 6), and a predictive model was built for group membership respectively. The model is based on a discriminant function based on linear combinations of the predictor variables that provide the best discrimination between the groups (Pallant, 2013). Cross-validation using the 'leave one out classification' produces more accurate and reliable outcomes, thus this was used in the discriminant analysis shown here.

The assumptions of discriminant analysis implies that each predictor variable is normally distributed, thus all the variables, i.e. *m/z* values, were transformed using the square root transformation, which demonstrated to be the best transformation to deliver a more

normal distribution. Following this, the data sets were submitted to standardization in order to get comparable variables as described in the equation below.

Equation 3.4

$$Standardised\ variable\ ({}^{m}/_{Z}) = \frac{\left(\sqrt{{}^{m}/_{Z}} - Median\right)}{Standard\ deviation}$$

3. 6 Sensitivity and Specificity

Sensitivity and specificity are used to evaluate a clinical test. Ideally, sensitivity correctly identifies all patients with the disease, i.e. true positives, and similarly specificity correctly identifies all patients who are disease free, i.e. true negatives (Lalkhen and McCluskey, 2008). Sensitivity and specificity are determined through the following equations given below (Lalkhen and McCluskey, 2008):

Equation 3.4

$$Sensitivity = \frac{true\ positives}{true\ positives + false\ negatives}$$

Equation 3.5

$$Specificity = \frac{true\ negatives}{true\ negatives + false\ positives}$$

The designation of the terms true positive, false positive, true negative and false negative are given below:

- ***True positive:*** the patient has the disease and the test is positive
- ***False positive:*** the patient does not have the disease but the test is positive
- ***True negative***: the patient does not have the disease and the test is negative
- ***False negative***: the patient has the disease but the test is negative

I will now discuss results of sets of experiments performed during my PhD studies.

Chapter 4

A potential method for comparing instrumental analysis of breath gas volatile organic compounds using standards calibrated in the gas phase

Considerable efforts have been undertaken to develop non-invasive diagnostic methods for detecting and monitoring disease through the analysis of volatile organic compounds (VOCs) which can potentially be used as biomarkers. These VOCs have been detected in exhaled breath, skin emanations, saliva, urine and faecal headspace (de Lacy Costello *et al.*, 2014). Breath analysis in particular has recently become a topic of interest for its potential to provide a non-invasive screening tool in early disease diagnosis (Lourenço and Turner, 2014). However such measurements have been limited by inconsistent evaluations of the concentrations of such VOCs by different instruments even when they are collected under apparently identical conditions. This shows that the lack of standardization between techniques is still a major challenge due to the vast disparity in the analytical tools employed, the sampling technique itself and the rich chemical diversity of exhaled breath at varied concentrations (Herbig and Beauchamp, 2014).

This chapter addresses a potential method for comparing the instrumental analysis of breath or headspace gas volatile organic compounds using three different analytical techniques, GC-MS coupled to thermal desorption (TD), SIFT-MS and PTR-MS through the use of standards calibrated for the gas-phase at physiologically representative concentrations. These analytical techniques have been widely used for medical research, particularly, within the breath analysis community in the identification and quantification of trace concentrations of VOCs in exhaled breath.

This work is currently under review in Journal of Breath Research.

4. 1 Product ions and branching ratios

A series of experiments were performed using standard aqueous solutions of acetone, ethanol, methanol, 1-propanol, 2-propanol and acetaldehyde were prepared according to Henry's law and calibrated for the gas-phase at 293 K ~ 20 ˚C as described in section 3.1.1.

Table 4.1 lists the ions observed and their branching ratios when these compounds were analysed by SIFT-MS and PTR-MS. These may be compared with earlier measurements discussed in chapter 3 (section 3.1.6).

Table 4. 1 - Comparison of the product ions identified by SIFT-MS and PTR-MS and their associated ion branching ratios (percentage in parentheses), respectively, calibrated for 5 ppm in a highly humid headspace, for a series of saturated alcohols (ROH), acetone (RO) and acetaldehyde (RO) present in human breath. Errors in the branching ratios are estimated to be less than 10%.

		SIFT-MS	***PTR-MS***
VOC	***Molecular formula (MW)***	***Product ions m/z and ion branching ratios (%)***	***Product ions m/z and ion branching ratios (%)***
Methanol	CH_3OH (32)	33 (41) ROH_2^+ 51 (28) $ROH_2^+ \cdot H_2O$ 69 (31) $ROH_2^+ \cdot (H_2O)_2$	33 (96) ROH_2^+ 51 (4) $ROH_2^+ \cdot H_2O$
Acetaldehyde	C_2H_4O (44)	45 (41) ROH^+ 63 (47) $ROH^+ \cdot H_2O$ 81 (12) $ROH^+ \cdot (H_2O)_2$	45 (54) ROH^+ 63 (46) $ROH^+ \cdot H_2O$
Ethanol	C_2H_5OH (46)	47 (15) ROH_2^+ 65 (45) $ROH_2^+ \cdot H_2O$ 83 (40) $ROH_2^+ \cdot (H_2O)_2$	45 (1) $[ROH\text{-}H_2]H^+$ 47 (51) ROH_2^+ 65 (48) $ROH_2^+ \cdot H_2O$
Acetone	C_3H_6O (58)	59 (39) ROH^+ 77 (61) $ROH^+ \cdot H_2O$	59 (100) ROH^+
1-Propanol	C_3H_7OH (60)	43 (70) R^+ 61 (2) ROH_2^+ 79 (6) $ROH_2^+ \cdot H_2O$ 97 (22) $ROH_2^+ \cdot (H_2O)_2$	43 (77) R^+ 61 (6) ROH_2^+ 79 (17) $ROH_2^+ \cdot H_2O$
2-Propanol	C_3H_7OH (60)	43 (54) R^+ 61 (5) ROH_2^+ 79 (29) $ROH_2^+ \cdot H_2O$ 97 (12) $ROH_2^+ \cdot (H_2O)_2$	43 (68) R^+ 59 (1) $[ROH\text{-}H_2]H^+$ 61 (12) ROH_2^+ 79 (19) $ROH_2^+ \cdot H_2O$

4. 2 Theoretical prediction of reaction rate coefficients

In order to compare these different methods and put the results in context, it is important to understand the theoretical prediction of reaction rate coefficients and thus determine headspace concentrations. A vast number of rate coefficients of VOCs with H_3O^+ at thermal energy have been experimentally determined using SIFT-MS with a reported accuracy better than 10% (Španěl *et al.*, 1995, Španěl *et al.*, 1997, Španěl and Smith, 1997, Smith and Španěl, 2011, Španěl *et al.*, 2002). The rate coefficients have been calculated using the parameterised trajectory formulation of Su and Chesnavich (Su and Chesnavich, 1982) within an estimated accuracy ± 10%. The rate coefficients at thermal conditions have been previously reported in the literature (Španěl and Smith, 1997, Španěl *et al.*, 1997).

In contrast the kinetics of ion-molecule reactions of PTR-MS are not well-defined, literature data on reaction rate coefficients is scarce, and can be sensitive to *E/N* (Tani *et al.*, 2003, Warneke *et al.*, 1996, Cappellin *et al.*, 2010). A standard value (2×10^{-9} cm^3 s^{-1}) for the proton transfer rate coefficient (k_{cap}) is often used to estimate trace-gas concentrations. Quantification is directly dependent on k_{cap}, which means that using the same proton transfer rate coefficient for different ion-molecule reactions, such an assumption may introduce a significant uncertainty to the quantification.

Proton transfer reactions between H_3O^+ and VOCs in both SIFT-MS and PTR-MS occur efficiently if the proton affinity of the acceptor molecule exceeds the proton affinity of the donor (H_2O), thus such reactions are exothermic. Exothermic proton transfer reactions in the gas-phase tend to be fast ($k \geq 10^{-9}$ cm^3 s^{-1}) and usually proceed at the collisional rate (Blake *et al.*, 2009, Bohme *et al.*, 1980).

An estimate of rate coefficients of exothermic ion-molecule reactions is possible, if the polarizability and the dipole moment of the reactant molecule are known. Therefore, in this study ion-neutral molecule collision theories have been applied, and the reaction rate coefficients were calculated according to the detailed description given elsewhere (Cappellin *et al.*, 2010). Table 4.2 shows the theoretical prediction of proton transfer rate coefficients, k_{cap} (T_{rot},KE_{cm}), for the proton transfer reactions occurring in the drift tube of PTR-MS at 373 K, calculated via Su and Chesnavich trajectory theory (Su, 1994). For comparison, reaction rate coefficients were calculated via the Langevin theory (Langevin, 1905), and Average Dipole Orientation (ADO) theory (Su and Bowers, 1973).

Table 4. 2 - Proton transfer reaction rate coefficients k_{cap} (T_{rot}, KE_{cm}) (10^{-9} cm^3 s^{-1}) between hydronium ion (H_3O^+) and selected VOCs at 373 K, T_{rot} is taken as the drift tube temperature. For comparison, reaction rate coefficients k_L, k_{ADO}, and k_{cap} (T_{eff}) are also reported and expressed in 10^{-9} cm^3 s^{-1}. Also given are their polarizabilities, α, expressed in units of 10^{-30} m^3 and their permanent dipole moment, μ, in Debye, D.

VOC	α (10^{-30} m^3)[a]	μ (D)[a]	k_L (10^{-9} cm^3 s^{-1})	k_{ADO} (10^{-9} cm^3 s^{-1})	k_{cap} (T_{eff}) (10^{-9} cm^3 s^{-1})	k_{cap} (T_{rot},KE_{cm}) (10^{-9} cm^3 s^{-1})
Acetone	6.40	2.88	1.57	3.10	2.24	3.10
Ethanol	5.26	1.69	1.46	2.14	1.89	2.36
Methanol	3.29	1.70	1.23	2.13	1.66	2.21
1-propanol	6.74	1.58	1.60	2.21	1.96	2.33
2-propanol	6.97	1.58	1.63	2.24	1.99	2.37
Acetaldehyde	4.59	2.75	1.38	2.97	1.93	2.80

[a] Reference (Lide, 2014).

For VOCs with oxygen in their chemical structure, the permanent dipole moment is much greater than 1.0 D and thus the permanent dipole moment contributes significantly to the rate coefficient. Calculations via the Langevin theory give rate coefficients, k_L, may be too low since the ion-dipole interaction is not accounted for. For instance, the Langevin theory rate coefficient for acetone is 1.57×10^{-9} cm^3 s^{-1}, while the corresponding value of the ADO theory is 3.10×10^{-9} cm^3 s^{-1}, which is greater than the Langevin theory by a factor of two. The rate coefficients, k_{ADO} , predicted from the ADO theory at 300 K tend to underestimate measured rate coefficients by typically by 10 – 20% (Su and Bowers,

1973). The rate coefficients k_L and k_{ADO} were calculated according to the equations 2.27 and 2.28 described in chapter 2.

Ion-molecule reactions in PTR-MS proceed with energies larger than thermal energy due to the existence of a homogenous electrical field. Increasing the electric field strength in the PTR-MS or decreasing the pressure or temperature of the drift tube in turn increases the drift velocity and, consequently, the kinetic energy of the ions. The temperature of the ions is higher than the drift-tube temperature as a result of the selective heating of the ions by the electric field. Therefore, this additional energy applied by the electric field should be taken into account. This means that is necessary to consider the centre-of-mass kinetic energy (KE_{cm}) for a collision between an ion and the neutral molecule (Viggiano and Morris, 1996, Ellis and Mayhew, 2014). Calculation of KE_{cm} is indicated in equation 2.31 described in chapter 2.

Trajectory calculations have been carried out by Su and Chesnavich (Su and Chesnavich, 1982), combining variational transition theory and classical trajectory analysis of ion-polar molecule capture collisions to derive rate coefficients, k_{cap}, for ion-polar molecule capture collisions. For a limited number of cases it has been shown that the effective temperature, T_{eff}, describes the internal energy of the ion reactants and depends only on the ratio E/N (Viggiano and Morris, 1996) and this is used in this work to determine rate coefficients for ion-molecule collisions, k_{cap}. Calculation of rate coefficients k_{cap} and effective temperature T_{eff}, are indicated in equation 2.29, equation 2.30, and equation 2.34 described in chapter 2.

Ions have three separate energy distributions, translational, rotational, and vibrational. It has been found that the average translational energy can be calculated from the drift velocity using the Wannier expression given in chapter 2, equation 2.32 (Wannier, 1953). As regards to vibrational energies, it has been reported in literature that in many cases vibrations do not affect the kinetics significantly, i.e. rate constants measured as a function of KE_{cm} at several temperatures do not vary significantly (Viggiano and Morris, 1996). The dependence of the rate constants upon rotational temperature of the neutral molecule is obtained by comparing data at a specific centre-of-mass kinetic energy at several different temperatures. Initially it was assumed that the rotational energy of the reactant ion would have a negligible effect on the rate constants but more recently it has been reported that the rotational energy of the neutral sometimes appears to have a significant effect on rate constants, leading to the conclusion that the rotational energy of the ions may affect rate coefficients in the same manner (Viggiano *et al.*, 1988). Theoretical and experimental results therefore indicate that rotational energy effects on ion-molecule rate coefficients must be considered, particularly for polyatomic ions (Viggiano *et al.*, 1988, Viggiano *et al.*, 1990, Viggiano and Morris, 1996). Su has developed a parameterization of the kinetic energy, giving capture rate constants, k_{cap} (T_{rot},KE_{cm}), within 5% error for ion-molecule systems with centre-of-mass kinetic energies ranging from thermal to several electron-volts (eV), and a wide range of temperatures (50-1000 K) (Su, 1994). Determination of the rate coefficient k_{cap} (T_{rot},KE_{cm}) is indicated in equation 2.35 described in chapter 2. Such a method is based on trajectory calculated results relating the Langevin rate constant to the capture rate constant by a factor, K_c, parameterised by KE_{cm} and rotational temperature (T_{rot}). The latest theory possibly provides the most accurate results, since the collision energy and rotational energy effects have been considered. In this work we have assumed the drift tube temperature

as rotational temperature (T_{rot}) assuming no effect of the electric field on the rotational energy of the neutral. The findings shown in Table 4.2 indicate that the rate coefficients k_{cap}, calculated at an effective ion temperature T_{eff}, underestimate the proton transfer rate coefficients by 10-40% compared to Su's rate coefficients k_{cap} (T_{rot},KE_{cm}), which is in agreement with earlier measurements by Cappellin *et al.* (Cappellin *et al.*, 2010). While the accuracy of the model to determine the rate coefficient is improved by utilising k_{cap}(T_{rot},KE_{cm}), the uncertainty is also increased compared to simpler models.

4. 3 Quantitative determination of VOC concentrations

4.3.1 Individual headspaces

Following on from the previous section, quantification of VOC concentrations may be accomplished if proton transfer reaction rate coefficients are known and Su's parameterised kinetic energy dependence of ion-molecule reaction rate coefficients k_{cap} (T_{rot},KE_{cm}) have therefore been used to determine the PTR-MS VOC concentration. The k_{cap} (T_{rot},KE_{cm}) calculations were drawn from equations 2.35 and 2.36 described in chapter 2.

PTR-MS VOC concentrations were derived in parts-per-billion by volume (ppbv) according to the detailed description given elsewhere (Cappellin *et al.*, 2012), and by using equations 2.42 and 2.43 presented in chapter 2 (taking into account the density of air in the reaction chamber). In summary, the method consists of using the proton transfer rate coefficient [k_{cap} (T_{rot},KE_{cm})] expressed in $cm^3\ s^{-1}$ between the hydronium ion (H_3O^+) and VOC (R), calculated previously; determining the reaction time (t) of the ions in the drift tube expressed in seconds, which is dependent on the drift tube length (L), the electric

potential (U) applied to the drift tube, and the mobility of the H_3O^+ ions in the drift tube which depend on the working conditions (pressure and temperature); and determining the counts per second (cps) of detected protonated parent ion (RH^+) and its fragment ions, and the counts per second (cps) of H_3O^+. The presence of H_3O^+ water clusters $[H_3O(H_2O)_n]^+$ ions, with n = 0, 1, 2 , 3 must also be considered. The mass-dependent transmission efficiency (*Tr*) of ions (H_3O^+ and RH^+) into the PTR-MS detection system was also taken into account since just a fraction of the total ion count reach the detector successfully.

The contribution of fragment ions and hydronium water-cluster ions was taken into account in the determination of VOC concentration via SIFT-MS and PTR-MS. In gas samples with higher humidity, such as breath gas, the number of water-cluster ions is significantly enhanced, and these may play a role as a reagent ion. The contribution of fragment ions for the determination of VOC concentration by PTR-MS was calculated according to equation 2.43 in chapter 2. The sum of the normalised counts per second for each fragment ion plus the protonated parent ion was made for the data acquired via SIFT-MS. Determination of VOC concentration by GC-MS was performed through the use of the internal standard approach, and calculated according to equation 3.3 in chapter 3.

Table 4.3 presents a comparison between the mass spectrometric techniques used and TD-GC-MS, for 5 ppm in the headspace, and this concentration was initially calculated according to Henry´s law at 20 °C. The ambient air temperature in the laboratory was measured simultaneously with experiments undertaken, and laboratory temperature was found to be higher than 20 °C. The temperature significantly influences VOC concentration in the headspace, therefore, concentrations were corrected for the

ambient air temperature in the laboratory at 20 °C. The compound-dependent sensitivity of SIFT-MS and PTR-MS is expressed in ncps/ppbv and is listed in Table 4.3.

Table 4. 3 - Determination of VOC concentrations calibrated to 5 ppm in the headspace at 20 °C and a comparison between the mass spectrometric techniques used and TD-GC-MS; SIFT-MS and PTR-MS sensitivity in ncps/ppbv within the range 10^1-10^3 ppbv.

VOC	***Concentration*** *(ppbv)*			***Retention time*** *(min)*	***Sensitivity*** *(ncps/ppbv)*	
	SIFT-MS [a]	***PTR-MS*** [b]	***TD-GC-MS*** [c]		***SIFT-MS*** [a]	***PTR-MS*** [b]
Acetone	4027	5089	5230	6.6	3.4	13.1
Ethanol	6167	3700	--	6.1	2.4	15.8
Methanol	6347	7203	--	4.2	2.0	7.8
1-propanol	4760	11816	3638	9.2	1.7	14.7
2-propanol	3189	10294	3848	7.1	1.9	12.4
Acetaldehyde	2600	1030	--	4.1	3.3	11.5

[a] Errors of ±20%, [b] ±50%, and [c] ±20% are assigned to the quantification via SIFT-MS, PTR-MS and TD-GC-MS, respectively. These were determined using the concentrations obtained via SIFT-MS, PTR-MS and TD-GC-MS.

Calibration curves for the six VOCs typically found in human breath are given within the range 10^1-10^3 ppbv and are represented in Figure 4.1 and Figure 4.2. The linearity of the response was verified for all compounds. Figure 4.1 indicates the calibration curves for the alcohols studied here and data acquired either using SIFT-MS *(a)* or PTR-MS *(b).* Figure 4.2 *(a)* indicates the calibration curve for acetone and Figure 4.2 *(b)* the calibration curve for acetaldehyde, either acquired via SIFT-MS or PTR-MS.

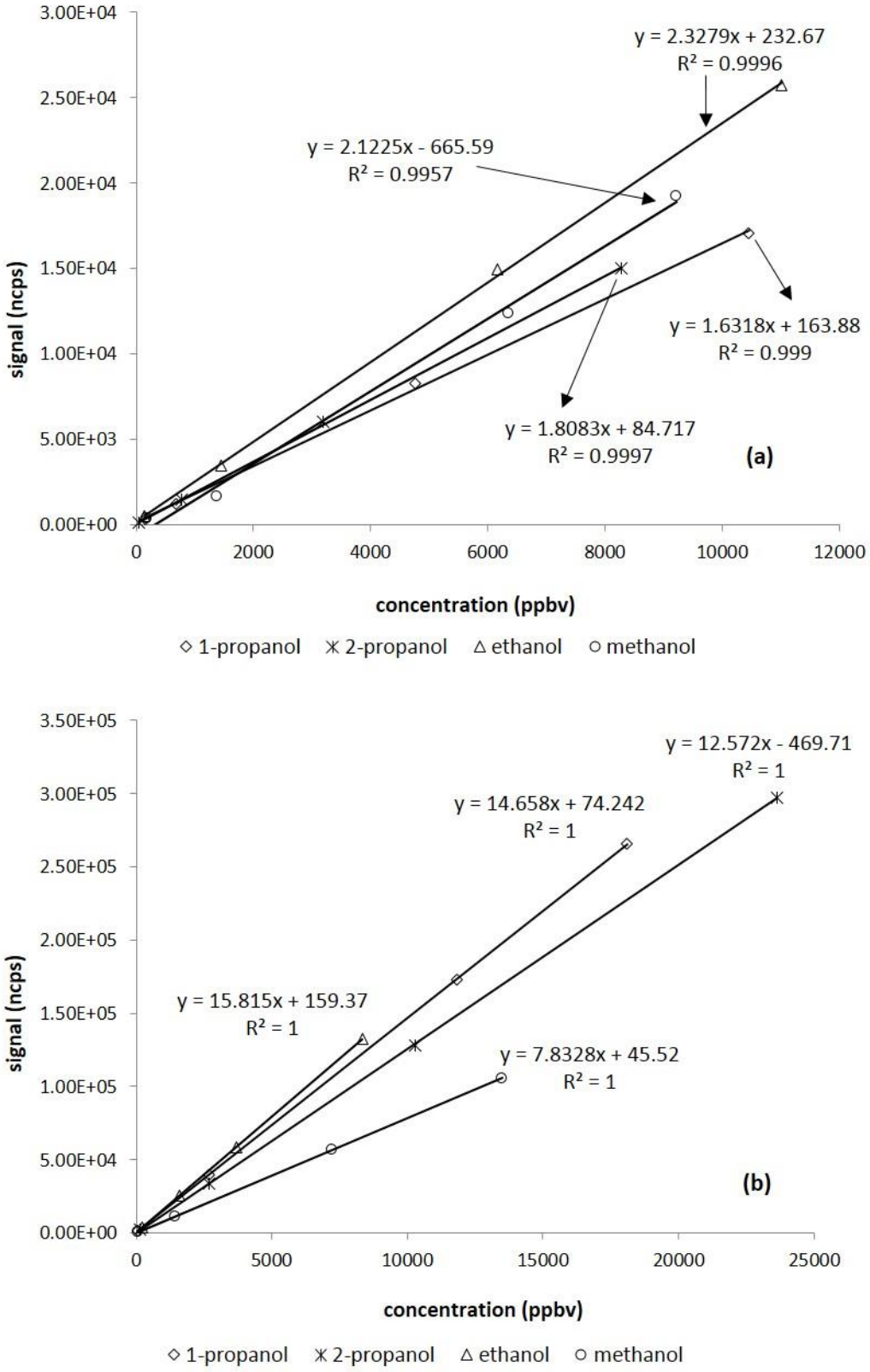


*Figure 4. 1 - Calibration curves for 1-propanol, 2-propanol, ethanol and methanol at physiologically representative concentrations; data acquired using SIFT-MS **(a)** and PTR-MS **(b)**.*

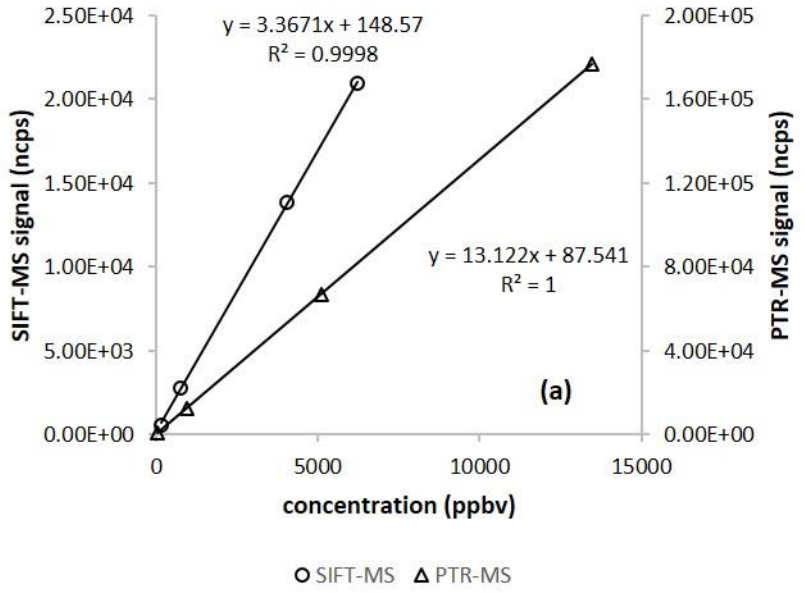
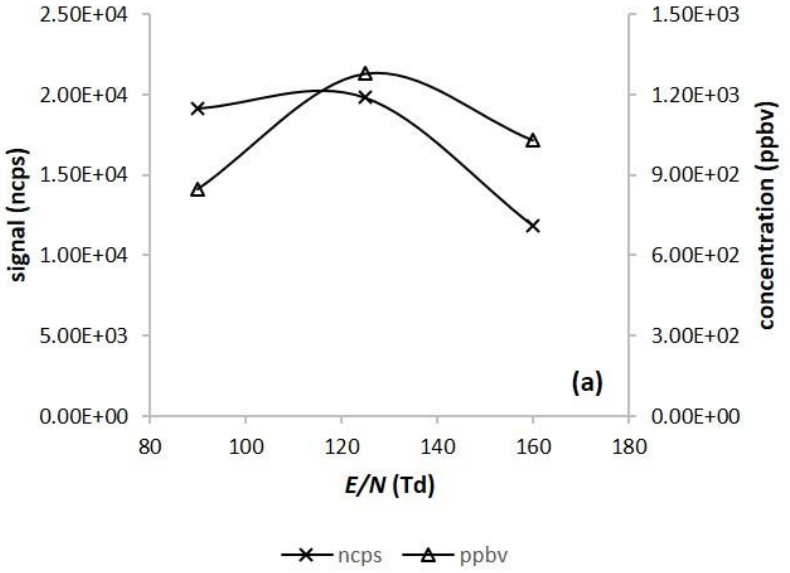
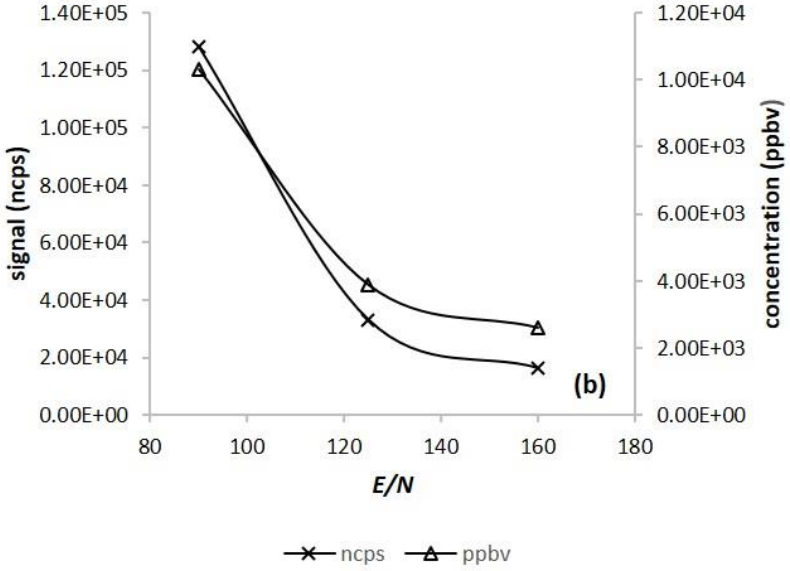


*Figure 4. 2 - Calibration plot for acetone **(a)** and acetaldehyde **(b)** at physiologically representative concentrations; data acquired using SIFT-MS and PTR-MS.*

Under standard operating conditions *E/N* should be in the range 120–130 Td – with the exception of the explosive 2,4,6 trinitrotoluene (TNT) (Sulzer *et al.*, 2012) whose sensitivity of detection is enhanced at higher *E/N* values – representing a compromise between reagent ion hydration on the one hand, and product ion fragmentation on the other (Lindinger *et al.*, 1998). This *E/N* range has therefore become the guideline for numerous experiments. Nevertheless, the effects on the underlying ion-chemistry must be carefully considered depending on the *E/N* ratio. The purpose of this study was not to analyse product ion distributions over *E/N* but our findings indicated that quantification via PTR-MS is sensitive to alterations in the reduced electric field (i.e. changes in the *E/N* ratio) and significantly influences the quantification of compounds in the gas-phase, as shown in Figure 4.3.

For most of the molecules used in this study product ion distributions as a function of *E/N* have been widely described in literature (Brown *et al.*, 2010). For instance, in acetaldehyde an increasing *E/N* leads to a significant fall in the signal (ncps) along with the decrease of the reaction time (μs), followed by the decrease of the reaction rate coefficient $k_{cap}(T_{rot}, KE_{cm})$. These dynamics are illustrated in Figure 4.3 presented below.





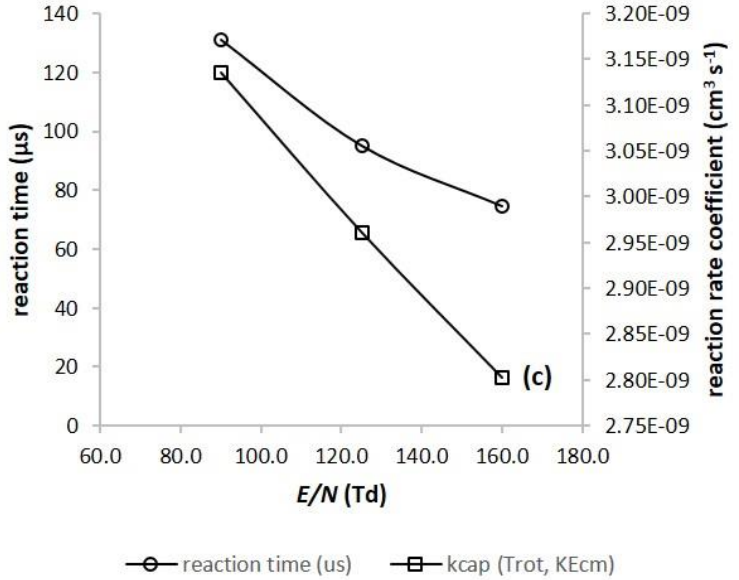


*Figure 4. 3 - **(a)** Concentration of acetaldehyde in ppbv and (**b**) 2-propanol in ppbv, calibrated to 5 ppm in the gas-phase, and normalised counts per second (ncps) as a function of reduced electric field strength (E/N = 90, 125, 160); **(c)** Acetaldehyde-H_3O^+ reaction time and reaction rate coefficient as a function of E/N = 90, 125, 160.*

4.3.2 TD-GC-MS results

The VOCs were trapped in stainless-steel TD sorbent tubes and further desorbed into the GC column. The MS detector was setup to scan from *m/z* 33.0 to *m/z* 260.0, therefore, methanol (*m/z* = 32) was not detected through this specific GC-MS method. In addition, the TD sorbent used is suited to the trapping of volatiles with at least three carbons in their molecular structure, which is why acetaldehyde and ethanol were difficult to trap in our apparatus. Calibration functions using the internal standard procedure for 1-propanol, 2-propanol and acetone at physiologically representative concentrations are represented in Figure 4.4.

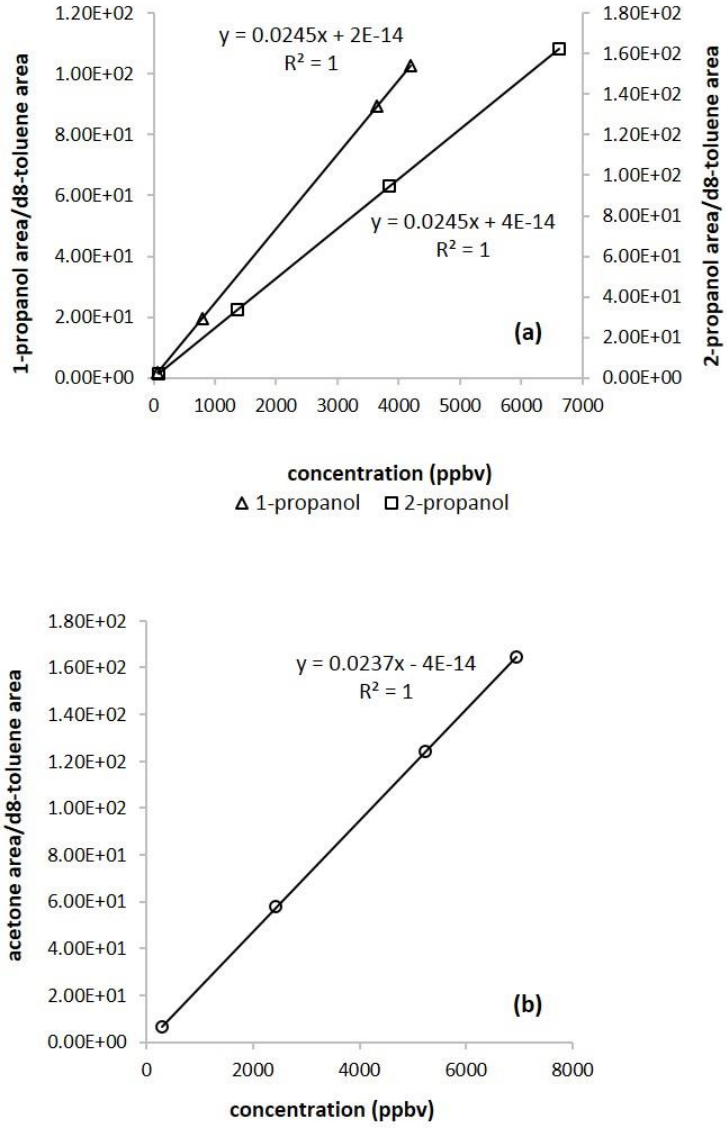


*Figure 4. 4 - Calibration functions using the internal standard procedure for 1-propanol, 2-propanol **(a)** and acetone **(b)**; data acquired using TD-GC-MS.*

The internal standard calibration curves plots the area of the analytes relative to the area of the internal standard (d8-toluene) against the concentration in the sample. The parameter determined relative to the internal standard is thus independent of deviations in the injection volume and possible variations in the performance of the detector (Hübschmann, 2009). These calibrations functions are specific for a sampling flow of 100 mL min^{-1} for 5 minutes, and 50 ng of internal standard d8-toluene was used.

Classic calibration in GC-MS systems frequently uses a calibration standard or a standard mixture at different concentrations (Hübschmann, 2009, Poole, 2012), however this method does not correct for any variation in the recovery of analytes by thermal desorption sampling techniques. The use of an internal standard improves accuracy and corrects for any variation in the recovery of analytes. In general, the internal standard approach is used to determine the concentration of an unknown sample investigated by GC-MS. The standard itself must not be present in the original sample and must be rapidly cleaned up from the column. Therefore, isotopically labelled standards should ideally be used (Hübschmann, 2009). For unknown VOC mixtures, the use of deuterated compounds enables rapid identification of the peak area which otherwise would be difficult to confirm the origin of the peak, either from the sample itself or the internal standard, avoiding misleading results. Thermal desorption procedures are generally calibrated using internal standard addition using deuterated toluene (Poole, 2012), where standard solutions are prepared and small volumes (typically 0.2 - 2 μL) loaded into each thermal desorption tube individually.

4.3.3 VOC mixtures

Since exhaled breath is a mixture of VOCs at different concentrations, which may respond differently once combined, the stability in the gas-phase of VOCs mixtures has been evaluated by using TD-GC-MS. VOC mixtures were prepared according to Henry´s law in the same manner as described in chapter 3 to give solutions of acetone, 1-propanol, 2-propanol, ethanol, acetaldehyde and methanol to give headspace concentrations of 0.1 ppm, 1 ppm, 5 ppm and 10 ppm. Internal standard calibration curves for the mixtures are given in Figure 4.5, where a sampling flow of 100 mL min^{-1} for 5 minutes was used and 50 ng of internal standard d8-toluene.

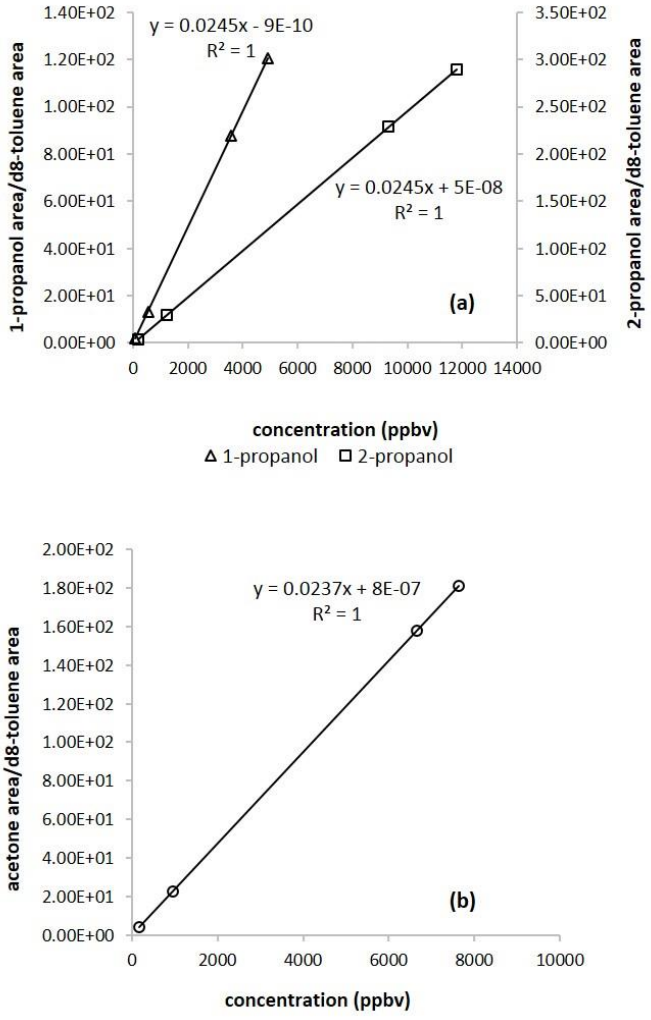


*Figure 4. 5 - Calibration functions using the internal standard procedure for 1-propanol and 2-propanol within a mixture of VOCs **(a)** and acetone **(b)**; data acquired using TD-GC-MS.*

Peaks elute along the GC column according to their affinity for the stationary phase of the column. Polar compounds interact weakly with low polarity stationary phases, as commonly used in VOCs analysis, and consequently, short retention times follow. It should be noted that competitive binding may occur and the significance of this for an accurate quantification is not yet fully understood. At a specific gas-phase concentration (i.e. 0.1 ppm, 1 ppm, 5ppm, and 10 ppm), the peak areas observed for the mixtures are significantly lower than for the individual solutions (Figure 4.6 (a), (b), (c)). This could possibly be due to competitive binding within the column. Among all the VOCs separated in the column, 2-propanol binds most strongly to the stationary phase due to its highest relative polarity and eluting later on at 7.1 minutes (Figure 4.6 (d)).

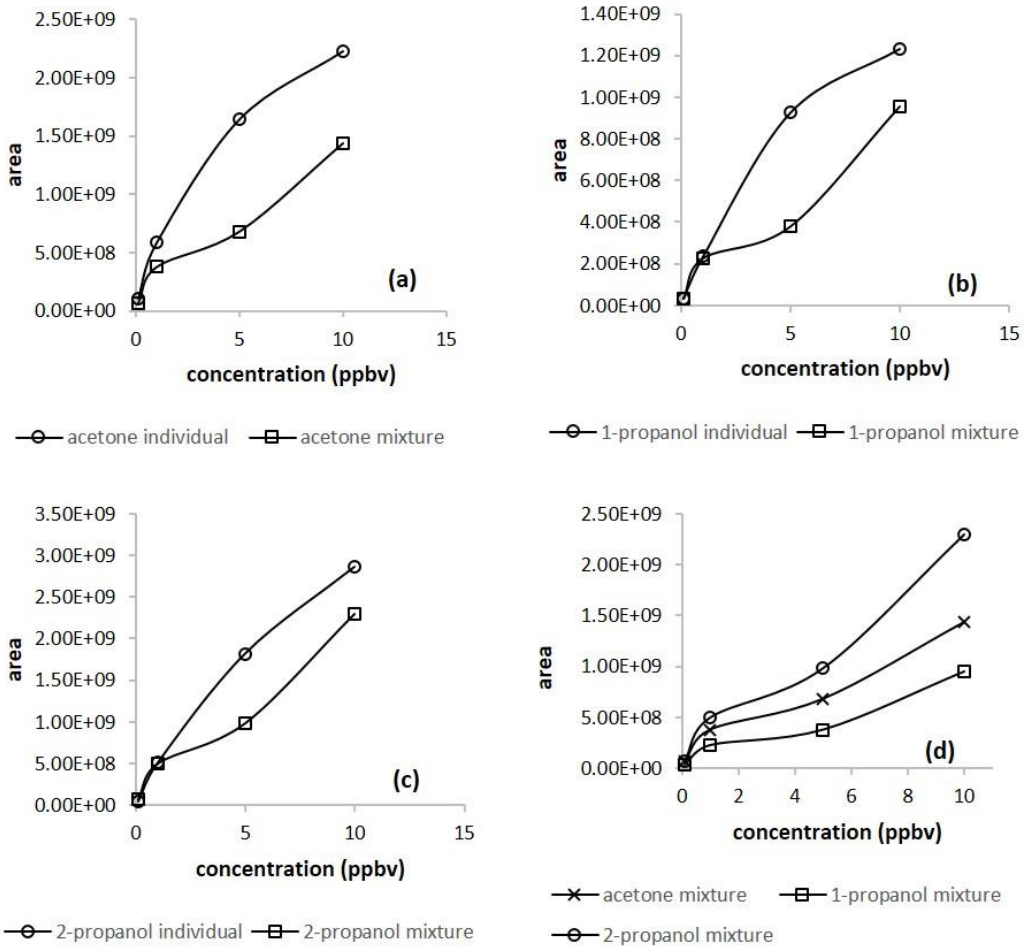


*Figure 4. 6 - Comparison of the peaks areas observed in chromatograms acquired using TD-GC-MS for individual calibration solutions, and peak areas acquired for a VOC mixture containing the six VOCs analysed **(a)** acetone, **(b)** 1-propanol, and **(c)** 2-propanol; **(d)** comparison of peak areas within the VOC mixtures at gas-phase concentrations of 0.1 ppm, 1 ppm, 5ppm, and 10 ppm.*

Gas chromatography–mass spectrometry has been recognised as the gold standard of analytical methodologies for many scientific tests. Its fundamental ability to effectively perform a qualitative analysis enables the identification of isomers within the sample which would be hard or nearly impossible to detect using a mass spectrometer only, particularly without separation of compounds present. However, issues remain as to the suitability of gas chromatography as a quantitative method of VOCs analysis, particularly if a TD system or solid-phase micro extraction (SPME) is used. Usually, the concentration of the substances of interest is too low for the direct measurement of an air sample and therefore enrichment on suitable adsorbents is necessary. In thermal desorption, the concentrated volatile components are desorbed by rapid heating of the adsorption tube, injected and stored in a cold trap, and subsequently, these are transferred to the GC column by rapidly heating the cold trap. This two-step desorption might have a crucial impact on the volatiles detected, i.e., competitive binding and desorption, not to mention thermal-lability of the VOCs.

The findings shown in this thesis indicated that most of the reactions result in multiple products ions and the abundance and stability of these ions strongly depends on the *E/N* ratio used.

Product ions and branching ratios were determined under specific working conditions and subsequently quantitative determination of VOC concentrations evaluated. Calibration curves determined using SIFT-MS, PTR-MS and TD-GC-MS are given within the range 10^1-10^3 ppbv.

This study demonstrates the potential way for standardising trace gas analysis of breath constituents given that different laboratories use different sampling and analytical techniques. In this work, the findings give slightly different results, but the approach taken of using standard headspaces enables a proper comparison to be made between three analytical techniques (SIFT-MS, PTR-MS and TD GC-MS) and enables these techniques to be tested for their accuracy and reproducibility.

In addition, this study demonstrates that using some techniques, such as GC-MS, VOCs with low molecular weight can be difficult to accurately quantify and some compounds are separated better than others, therefore this technique is only semi-quantitative.

Chapter 5

A longitudinal study of the VOC profile emitted by the faecal headspace of two different mouse models of diabetes

This chapter discusses the potential use of VOCs as a potential non-invasive method for monitoring the progress of diabetes. Mouse models of diabetes were used to identify potential indicators in VOCs emitted from faeces related to changes in blood glucose in type 2 diabetes (T2D). The results shown here were from studying two different mouse models of diabetes, namely, the Cushing´s syndrome mouse model of type 2 diabetes and single *Afmid* knockout mice exhibiting impaired glucose tolerance, and these were compared to their wild-type (WT) littermates respectively.

This animal work is under preparation for publication.

5. 1 A longitudinal study of the VOC profile emitted by the faecal headspace of Cushing´s syndrome mouse model of type 2 diabetes

As discussed in chapter 1, Cushing´s mice are prone to developing type 2 diabetes, i.e. hyperglycaemia and insulin resistance (Bentley *et al.*, 2014). In the present study the disease was tracked across the four time points (8, 12, 16 and 20 weeks of age) for the mutant (het) Cushing's mice and the wild-type (WT) littermates. A cohort of 65 mice (21 males and 44 females) was used of whom 30 were genotyped as mutants and 35 WT littermates.

The body weight of the animals is normally distributed, therefore a T-test was used to compare the mean score of weights between mutants (het) and WT littermates. The mutation had statistically significant effects on body weight (Figure 5.1). The mutant (het) animals were significantly heavier ($p < 0.05$) than the WT littermates.

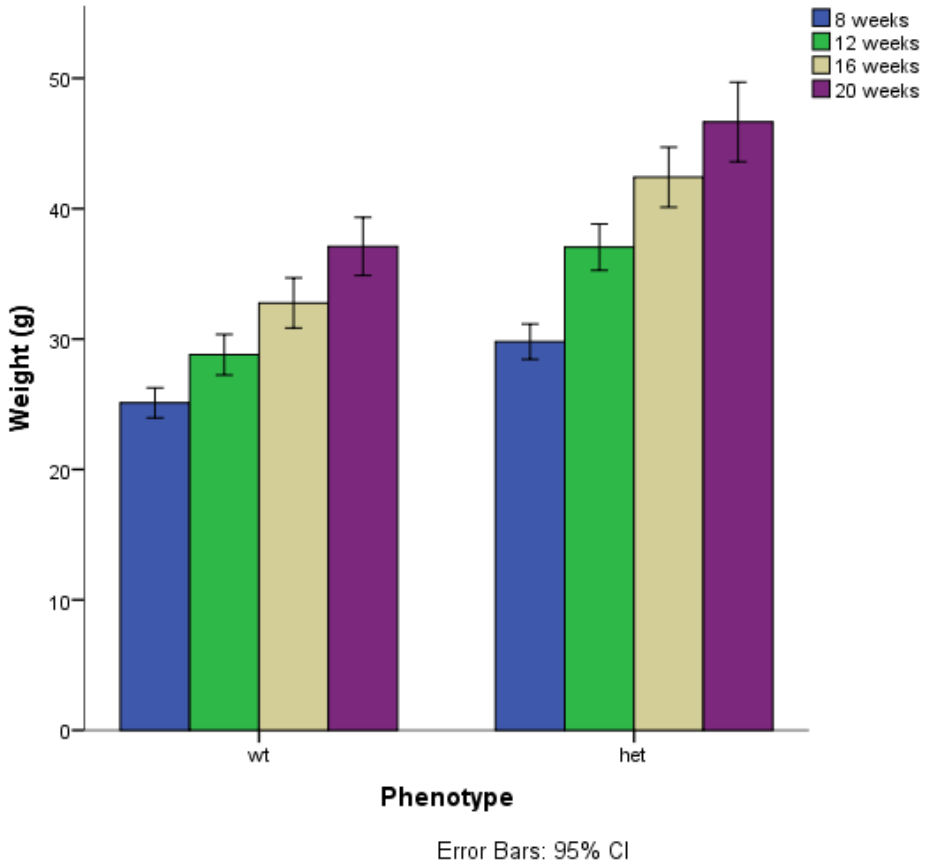


Figure 5. 1 – Body weights along the age of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using T- test ($p < 0.05$).

One-way repeated measures ANOVA with Bonferroni post-hoc test was conducted to compare the animal's weight along the age. The results indicated that the body weight of the animals significantly change over the four months of study. Thus, there was a statistically significant effect across the four periods of age, Wilk's Lambda = 0.155, $p < 0.05$.

Female mutant (het) mice were significantly heavier from 16 weeks of age on a B6-C3PDE background (Figure 5.2).

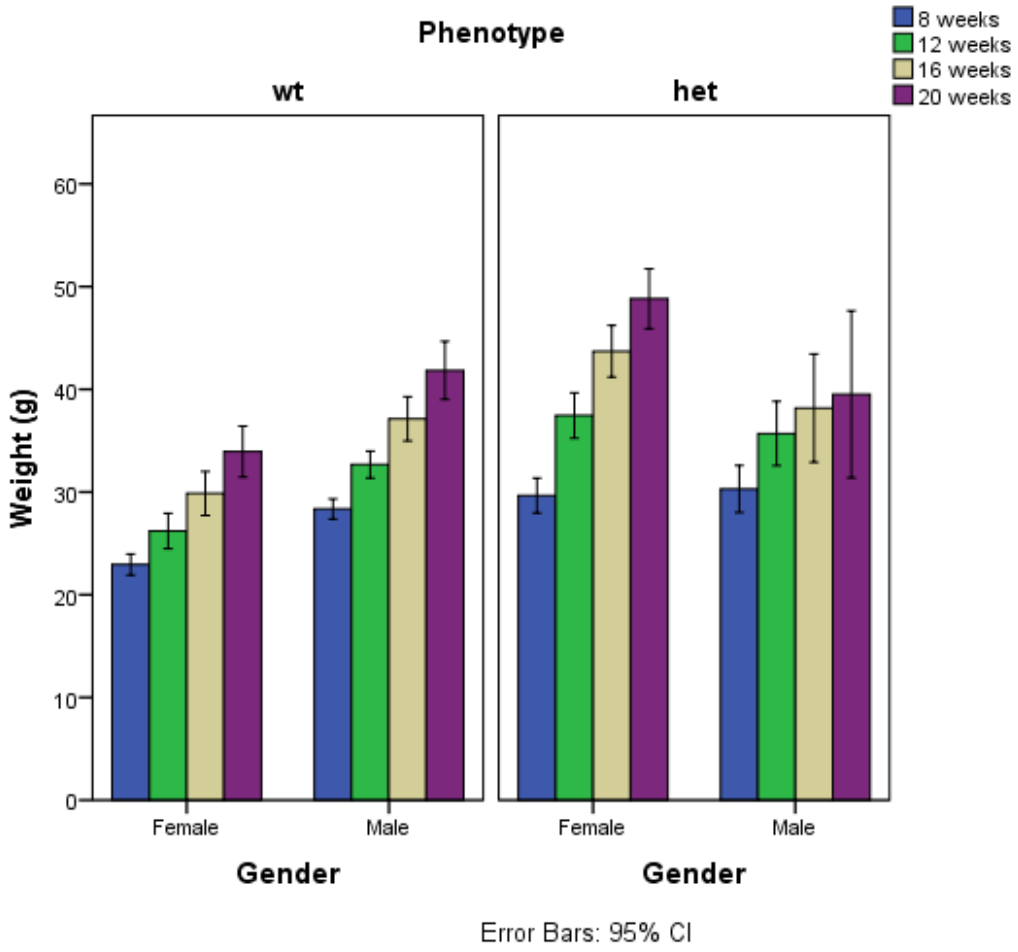


Figure 5. 2 – Body weights along the age of male and female Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using one-way repeated measures ANOVA with Bonferroni post-test.

Blood glucose was taken for unfasted animals with the blood being sampled at the same period of the day using an AlphaTRAK2 blood glucose meter. The performance and accuracy of the AlphaTRAK2 blood glucose meter was previously evaluated and the average coefficient of variation (mouse) determined is 5.9% (AlphaTRAK2). The mutation had statistically significant effects on blood glucose (Figure 5.3). The data indicate that male and female mutant mice had hyperglycaemia (blood glucose > 11 mmol L^{-1}) by the age of 16 weeks (Figure 5.3). In addition, the mutant male mice developed significant increased concentrations of blood glucose ($p < 0.05$) from 16 weeks of age compared to the WT littermates (Figure 5.4). Blood glucose sampling in mice induces stress in the animals and blood glucose variability might arise from that fact. In addition, some mice are more susceptible to obesity and diabetes than others, and insulin resistance

manifested at different time points. The estrous cycle in females significantly inserts some variability into the data.

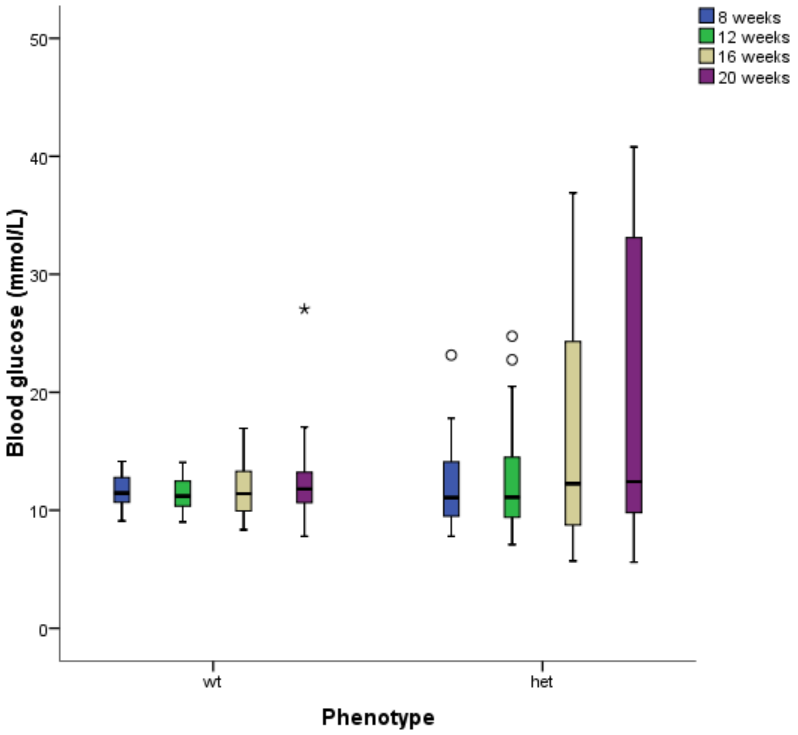


Figure 5. 3 – Blood glucose concentrations along the age in Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using T- test ($p < 0.05$).

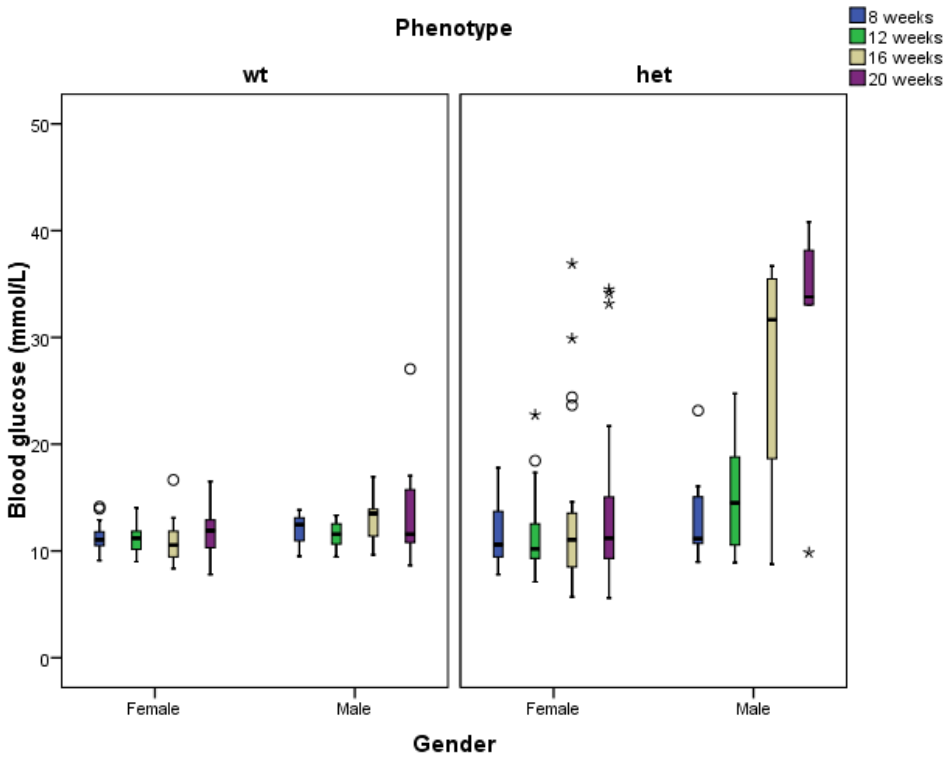


Figure 5. 4 – Blood glucose concentrations along the age of male and female Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Data analysed using Mann Whitney's test ($p < 0.05$).

Blood collection was undertaken after terminal anaesthesia of the animals for further analysis of insulin plasma concentrations. Therefore, the plasma insulin concentrations indicated in Figure 5.5 do not correspond to 20 weeks of age but a later stage once the animals were culled, i.e. animals born in April and May were culled at 28 weeks of age, and animals born in June were culled at 24 weeks of age. The insulin concentration of three animals are above the upper limit for the ELISA assay thus a value of 15 mmol L^{-1} was assigned for these animals.

Mutant animals (het) had statistically significant elevated plasma insulin concentrations *(median = 1.2 × 10^1, n = 30)* after 20 weeks of age once compared to WT littermates *(median = 2.3, n = 35)*, $p < 0.05$ (Figure 5.5). Insulin secretion is triggered by elevated glucose concentration in the blood, supporting the hypothesis that Cushing's mice develop a diabetic phenotype. There is no significant difference between insulin levels of males and females.

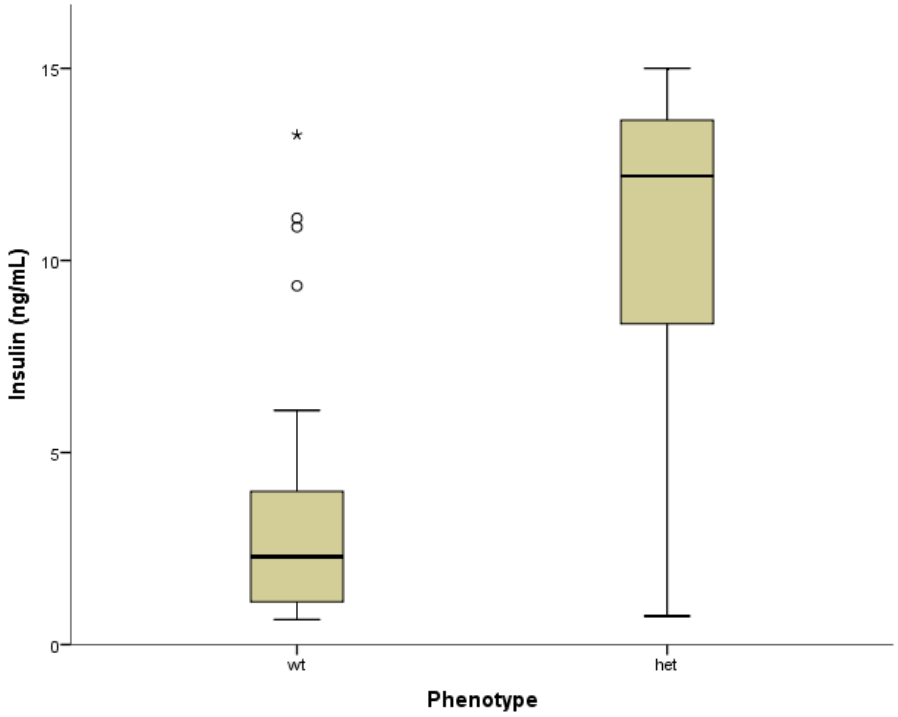


Figure 5. 5 – Plasma insulin concentrations of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background. Plasma insulin concentrations were taken once the animals were culled just after 20 weeks of age. Data analysed using Mann Whitney's test (p < 0.05)

5.1.1 Univariate Analysis using SIFT-MS data

The faecal headspace profile of Cushing´s mice was acquired using SIFT-MS at The Open University, in the form of *m/z* versus counts per second. The Kolmogorov-Smirnov statistic and Shapiro-Wilk statistic assessed the normality of the distribution of scores along the data set. Results indicated that the data are not normally distributed for H_3O^+, NO^+ and O_2^+ precursor ions. Accordingly, a non-parametric statistic technique, Mann-Whitney U Test, was applied to test the differences between the two independent groups of variables, i.e. mutant (het) mice and WT littermates, and the same test was applied at each time point. The Mann-Whitney U test revealed significant differences in the VOCs levels of heterozygous mutant (het) animals and WT littermates for H_3O^+, NO^+ and O_2^+ precursor ions.

SIFT-MS H_3O^+ data were obtained from the headspace of faecal samples from a cohort of 65 mice (21 males and 44 females) of whom 30 were genotyped as mutants and 35 WT littermates. Using H_3O^+ as the precursor ion, statistically significant differences in the VOC levels of faecal headspace from mutant (het) mice and WT littermates were identified and given in Table 5.1.

Figure 5.6 shows an illustrative mass spectrum zoomed in of one faecal headspace sample from a mutant (het) Cushing´s mouse. Data were acquired by SIFT-MS H_3O^+, *m/z* range 10-140, acquisition time of 5 seconds and six iterations.

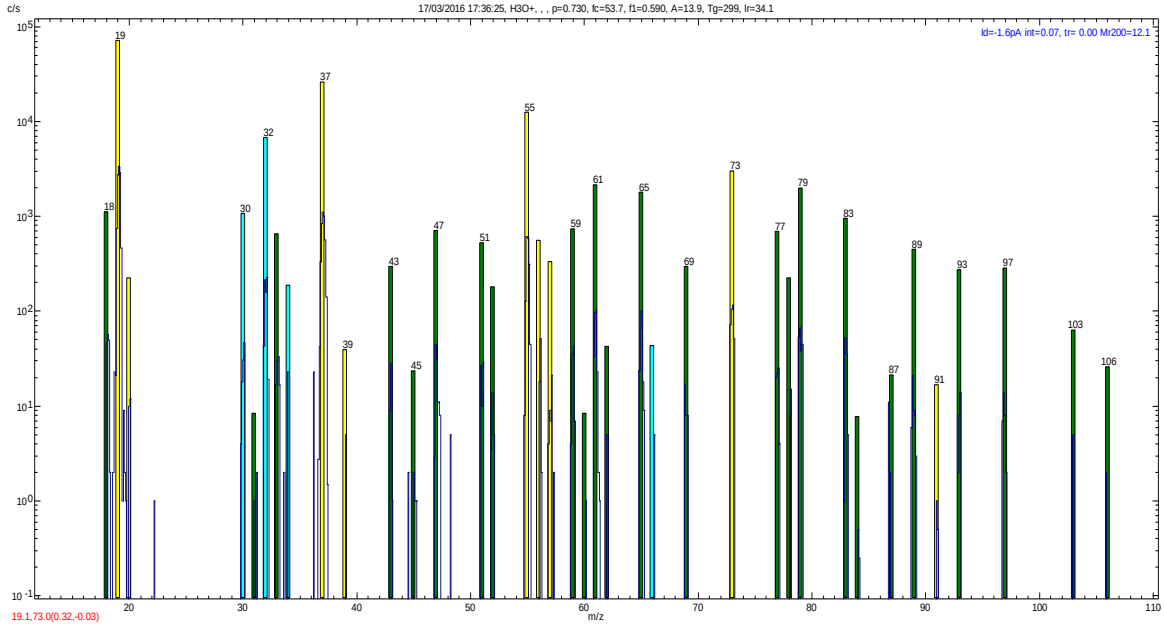


Figure 5. 6 – Illustrative mass spectrum of a faecal headspace sample from a mutant (het) Cushing´s mouse (animal ID 4.1h). Data acquired by SIFT-MS H_3O^+, m/z range 10-140, acquisition time of 5 seconds and six iterations.

Table 5.1 indicates the median values for statistically significant *m/z* values tested over the age (8, 12, 16 and 20 weeks of age), considering the full cohort of WT littermates and mutants respectively (n = 35, n = 30). The data were analysed using Mann Whitney´s test.

Table 5. 1 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of Cushing´s mice faecal headspace and its significance levels acquired using H_3O^+ precursor ion. Data were analysed using Mann Whitney´s test ($p < 0.05$).

			Median		
		Possible compound (s)	***wt***	***het***	
	Ion		***n = 35***	***n = 30***	*Significance level (p)*
8 weeks	*m/z* 57	Propanoic acid	1.1×10^2	3.4×10^2	0.014
	m/z 61	Acetic acid [a]	1.7×10^3	5.7×10^3	0.047
	m/z 65	Ethanol	1.3×10^3	2.0×10^3	0.042
	m/z 77	Acetone	3.3×10^2	5.7×10^2	0.023
	m/z 79	Acetic acid	9.8×10^2	2.5×10^3	0.001
	m/z 93	Propanoic acid	5.5×10^1	5.4×10^2	0.031
	m/z 97	Acetic acid	0.0	5.1×10^2	0.003
	m/z 107	Butanoic acid	9.3×10^1	4.3×10^2	0.030
12 weeks	*m/z* 83	Ethanol	8.9×10^2	2.7×10^2	0.017
16 weeks	*m/z* 33	Methanol	4.8×10^3	3.2×10^3	0.035
	m/z 38	Unknown	2.9×10^2	3.1×10^1	0.007
	m/z 89	Butanoic acid	3.3×10^3	1.1×10^3	0.043
20 weeks	*m/z* 43	Propanol	3.8×10^2	9.5×10^2	0.008
	m/z 45	Acetaldehyde	6.8×10^2	1.7×10^2	0.009
	m/z 47	Ethanol	9.3×10^2	2.2×10^3	0.001
	m/z 57	Propanoic acid	2.1×10^2	5.5×10^2	0.008
	m/z 61	Acetic acid	2.6×10^3	7.6×10^3	0.029
	m/z 65	Ethanol	2.1×10^3	6.5×10^3	0.000
	m/z 77	Acetone	3.1×10^2	1.8×10^3	0.004
	m/z 79	Acetic acid	2.0×10^3	4.3×10^3	0.010
	m/z 83	Ethanol	1.2×10^3	2.5×10^3	0.028
	m/z 93	Propanoic acid	8.1	4.5×10^2	0.002
	m/z 97	Acetic acid	0.0	1.0×10^3	0.001

[a] It is best to use NO^+ *m/z* 90 for acetic acid identification and *m/z* 43 for propanol identification

The *m/z* values 57 and *m/z* 93 are likely to be propanoic acid and these are shown to be statistically significant at 8 and 20 weeks of age. Although butanol might overlap at *m/z* 57 and *m/z* 93 potentially arising from the butyrate metabolism within the gut, i.e. reduction of butanoic acid to 1-butanol (Dash *et al.*, 2014, Garner *et al.*, 2007). Boxplots are indicated in Figure 5.7.

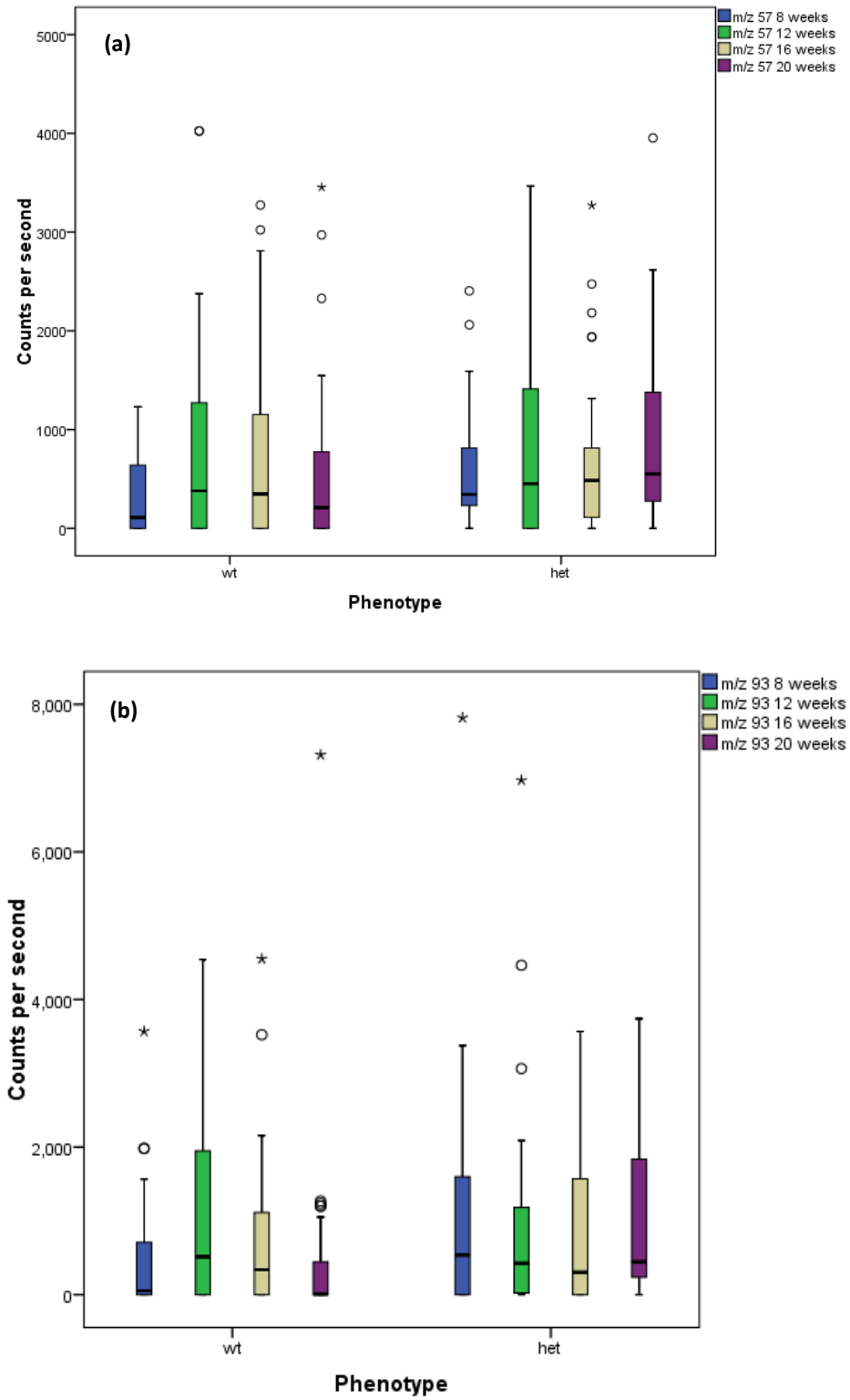


Figure 5. 7 – Boxplots of m/z 57 ***(a)****, m/z 93* ***(b)****; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).*

For H_3O^+ data set, *m/z* values 61, *m/z* 79 and *m/z* 97 revealed significant increased counts per second for mutant (het) mice compared to WT littermates, statistically significant at 8 and 20 weeks of age. These *m/z* values may be attributed to acetic acid because the NO^+ data set revealed *m/z* 90 to be statistically significant (as discussed in the section below). Boxplots are given in Figure 5.8.

The *m/z* values 47, *m/z* 65 and *m/z* 83 are likely attributed to ethanol and these were significantly higher at 20 weeks of age in the faecal headspace of the mutant (het) mice. Ethanol might result from the butyrate metabolism within the gut, converting acetic acid to acetaldehyde which in turn is converted to ethanol (Dash *et al.*, 2014). At 8 weeks of age only *m/z* 65 was statistically significant; and *m/z* 83 significant at 12 weeks of age in which WT littermates had significantly higher levels of ethanol. Boxplots are indicated in Figure 5.9.

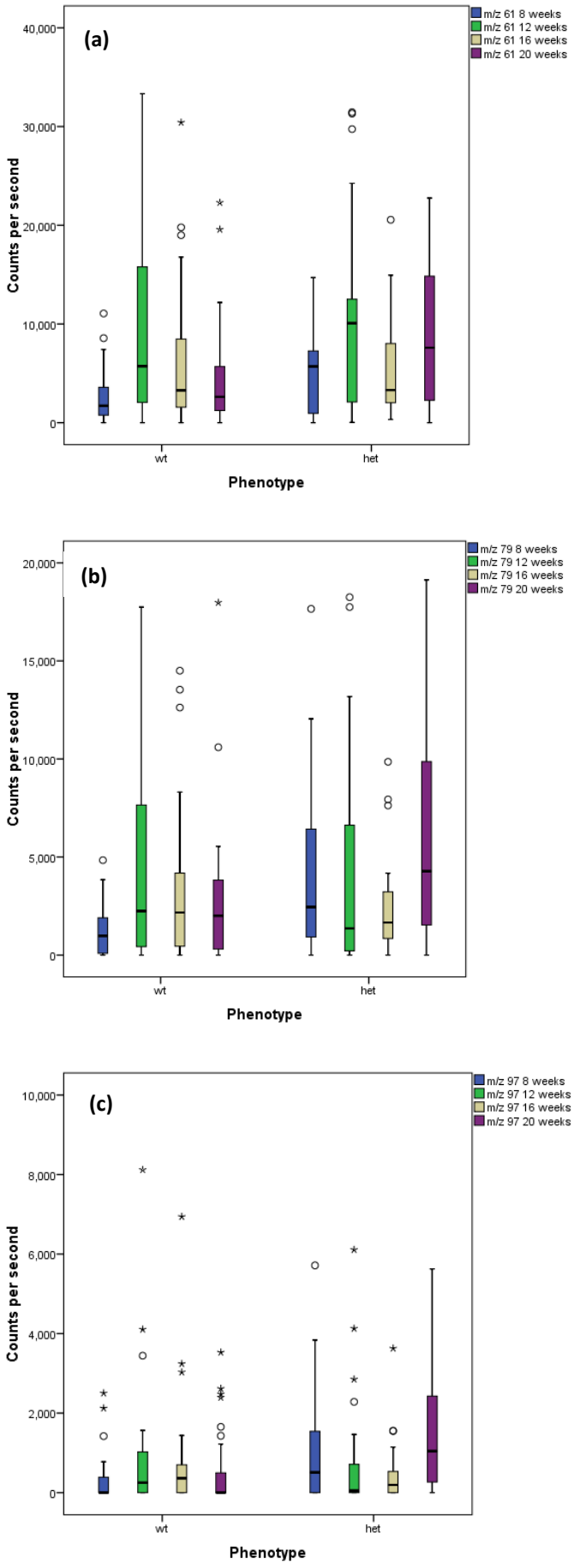


Figure 5. 8 – Boxplots of m/z 61 ***(a),*** *m/z 79* ***(b)****, m/z 97* ***(c)*** *are likely to be acetic acid; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).*

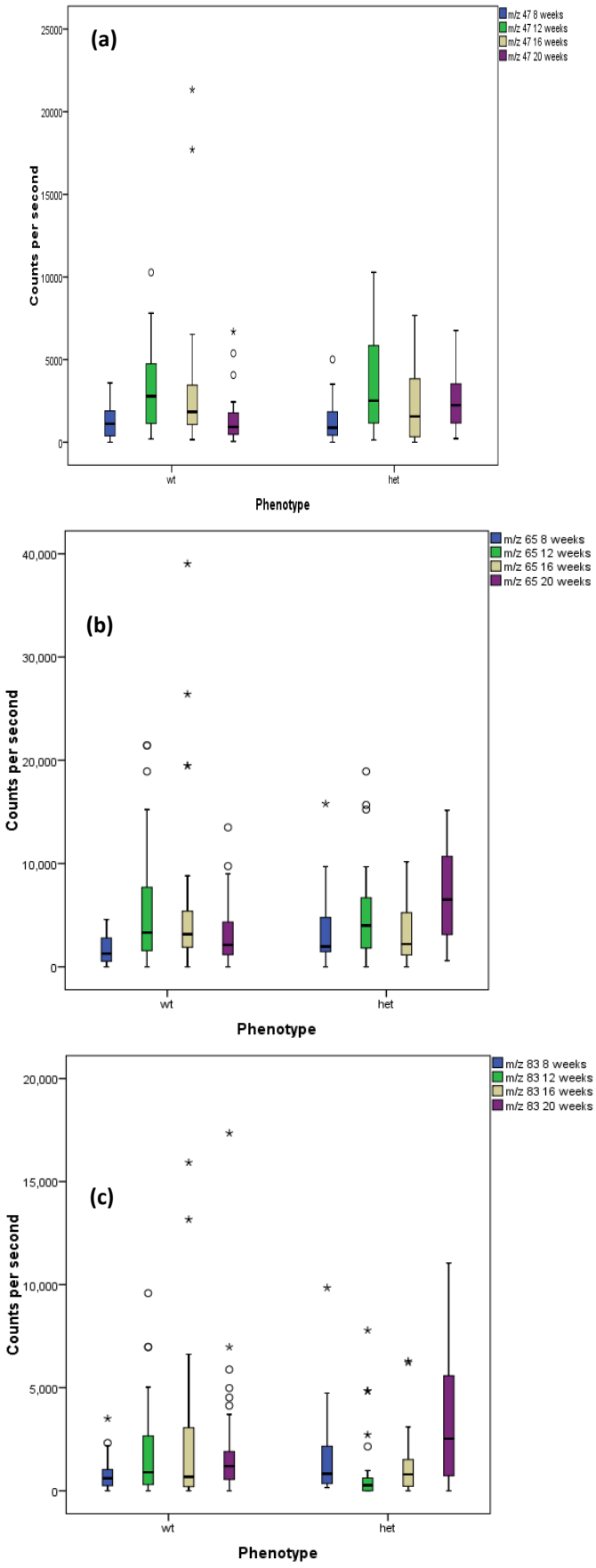


Figure 5. 9 – Boxplots of m/z 47 ***(a),*** *m/z 65* ***(b)****, m/z 83* ***(c)*** *are likely to be ethanol; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).*

Statistically significant higher counts per second of *m/z* 77 were also observed in mutant (het) mice faecal headspace at 8 and 20 weeks of age, which is most likely acetone (Figure 5.10).

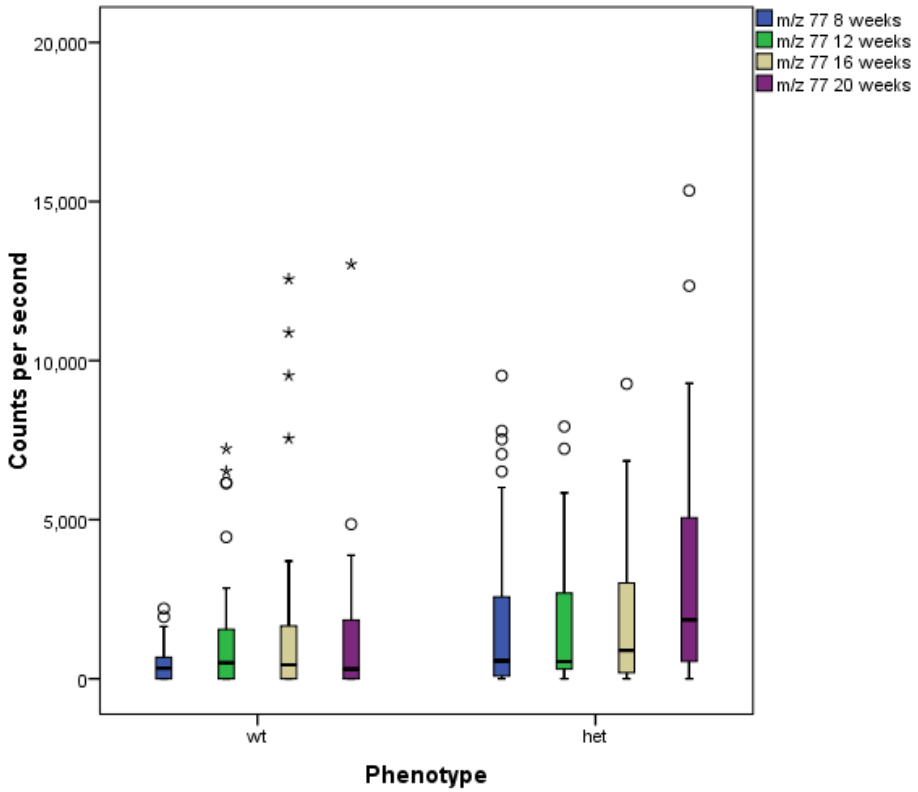


Figure 5. 10 – Boxplots of m/z 77; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).

For *m/z* 89 and *m/z* 107, which is tentatively identified as butanoic acid (although it could be many other compounds, i.e. pentanol, benzaldehyde and ethyl acetate) was detected in significant higher levels in mutant (het) mice at 8 weeks of age (Figure 5.11).

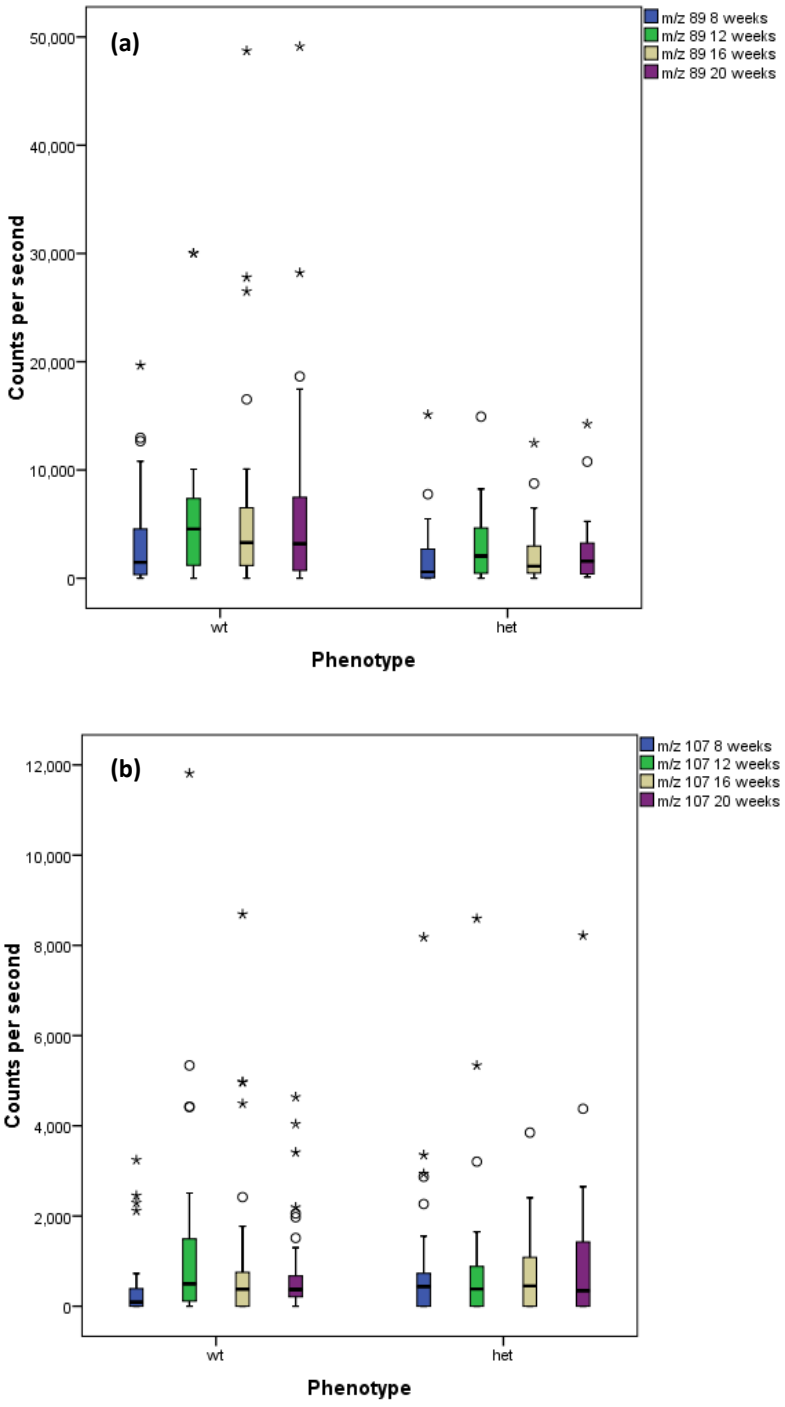


Figure 5. 11 – Boxplots of m/z 89 ***(a),*** *m/z 107* ***(b)****; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS* H_3O^+ *and analysed using Mann Whitney's test ($p < 0.05$).*

m/z 43, tentatively identified as propanol (which also produces ions at *m/z* 61 and *m/z* 79) and *m/z* 43 ion was detected in significant higher levels of faecal headspace of mutant (het) mice at 20 weeks of age (Figure 5.12).

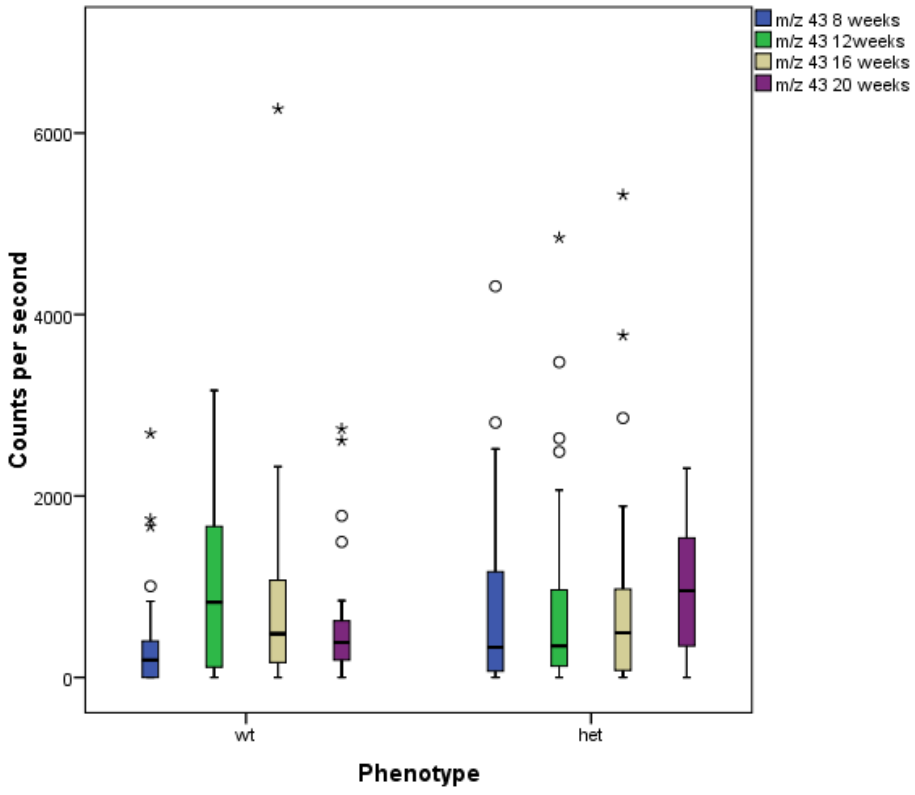


Figure 5. 12 – Boxplots of m/z 43; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).

Correlations and age effect in faecal headspace profile

The relationship between counts of *m/z* values and blood glucose levels was investigated using the non-parametric Spearman's rho correlation coefficient. Outliers have not been removed throughout this study. Three medium correlations ($0.30 \leq r \leq 0.49$) were found between specific *m/z* values and blood glucose levels respectively. Thus, medium correlation ($0.30 \leq r \leq 0.49$) for *m/z* 38 at 20 weeks of age ($r = 0.338$, $n = 65$, $p = 0.006$); *m/z* 93 at 16 weeks of age ($r = 0.298$, $n = 65$, $p = 0.016$); and *m/z* 59 at 16 weeks of age ($r = 0.251$, $n = 65$, $p = 0.044$).

Friedman test (repeated measures) suggest that are significant differences in *m/z* 61 (likely acetic acid according to NO^+ results) across the age ($p < 0.001$). A Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a statistically significant increase of acetic acid with age, i.e. from 8 to 12 weeks ($z = -4.048$, $p < 0.001$). In addition, a medium correlation was found between *m/z* 97 (acetic acid) and blood glucose levels at 20 weeks of age ($r = 0.300$, $n = 65$, $p = 0.015$).

A Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a statistically significant increase in *m/z* 65 (possibly ethanol) between 8 to 12 weeks of age ($z = -4.153$, $p = 0.000$); and between 8 to 20 weeks of age ($z = -3.728$, $p = 0.000$).
In addition, a negative correlation ($r = -0.302$, $n = 65$, $p = 0.014$) was found between *m/z* 83 and blood glucose levels at 8 weeks of age; small correlation for *m/z* 83 at 16 weeks of age ($r = 0.278$, $n = 65$, $p = 0.025$); and small correlation for *m/z* 83 at 20 weeks of age ($r = 0.244$, $n = 65$, $p = 0.050$).

SIFT-MS NO^+ data set: the Mann-Whitney U test revealed statistically significant differences in the VOCs levels of mutant (het) mice and WT littermates, were identified and given in Table 5.2.

Table 5. 2 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of Cushing´s mice faecal headspace and its significance levels acquired using NO^+ precursor ion. Data analysed using Mann Whitney´s test (p < 0.05).

			Median		
		Possible compound (s)	***wt***	***het***	
	Ion		***n = 35***	***n = 30***	*Significance level (p)*
8 weeks	*m/z* 71	Butanoic acid	0.0	1.2×10^2	0.036
	m/z 77	Propanol	0.0	3.1×10^2	0.005
	m/z 90	Acetic acid	2.4×10^2	6.8×10^2	0.017
	m/z 104	Propanoic acid	0.0	2.2×10^2	0.000
	Ion	***Possible compound (s)***	***wt***	***het***	***Significance level (p)***
12 weeks			***n = 23***	***n = 24***	
	m/z 104	Propanoic acid	0.0	3.1×10^2	0.014
	Ion	***Possible compound (s)***	***wt***	***het***	***Significance level (p)***
16 weeks			***n = 35***	***n = 30***	
	---	---	---	---	----
	Ion	***Possible compound (s)***	***wt***	***het***	***Significance level (p)***
			n = 35	***n = 30***	
20 weeks	*m/z* 77	Propanol	0.0	1.9×10^2	0.011
	m/z 88	Acetone	3.1×10^2	1.1×10^3	0.013
	m/z 104	Propanoic acid	0.0	4.5×10^2	0.003

In SIFT-MS NO^+ data set revealed that *m/z* values 71, 88, 90, and 104 are statistically significant, supporting the idea that these ions are in fact butanoic acid, acetone, acetic acid and propanoic acid respectively. Boxplots are represented below in Figures 5.13, 5.14, 5.15 and 5.16.

In addition, a medium correlation (r = 0.319, n = 47, p = 0.029) was found between *m/z* 90 at 16 weeks of age and blood glucose levels.

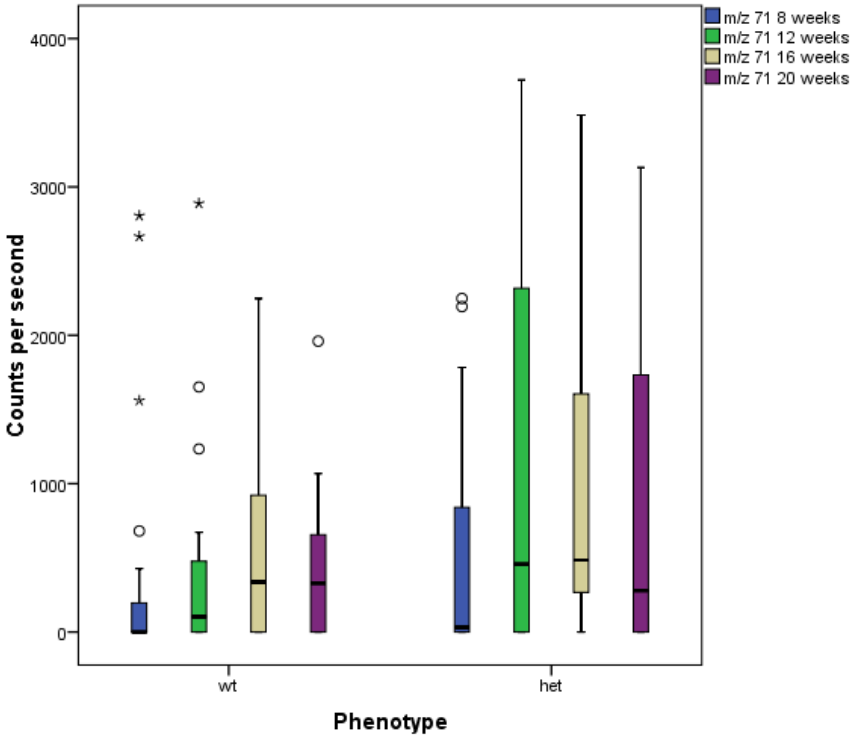


Figure 5. 13 – Boxplots of m/z 71; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

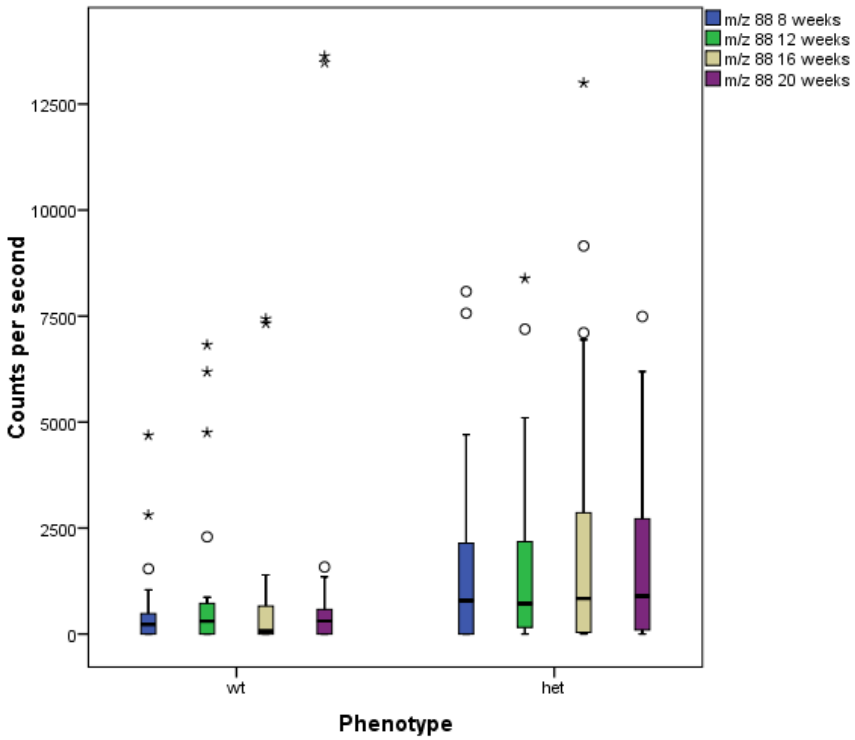


Figure 5. 14 – Boxplots of m/z 88; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

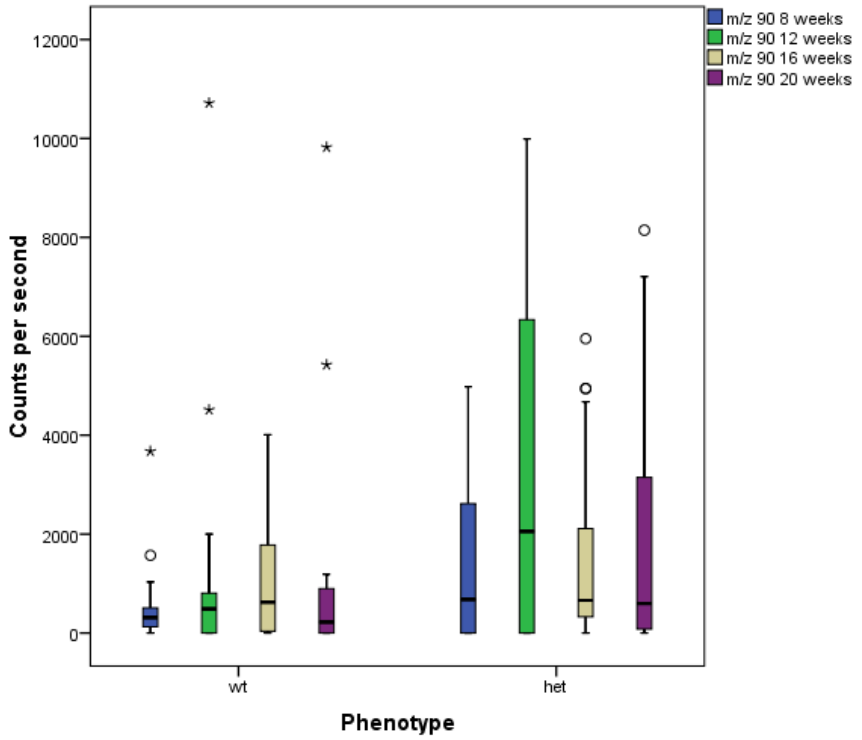


Figure 5. 15 – Boxplots of m/z 90; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

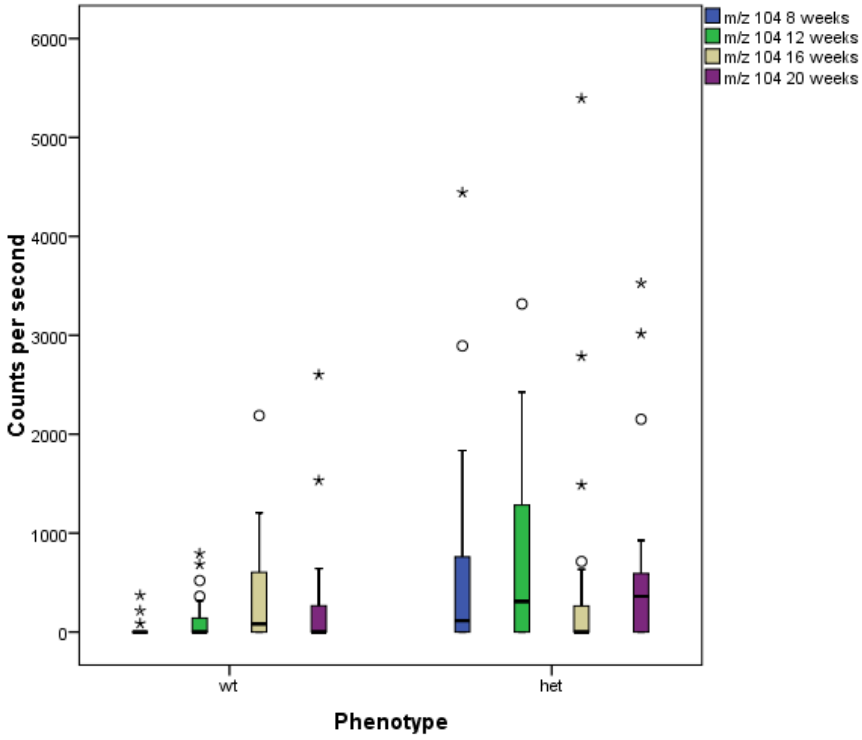


Figure 5. 16 – Boxplots of m/z 104; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

m/z 77 was shown to be statistically significant at 8 and 20 weeks of age, and this could be propanol which yields *m/z* 59 through hydride transfer and *m/z* 77 with association of one water molecule (Figure 5.17).

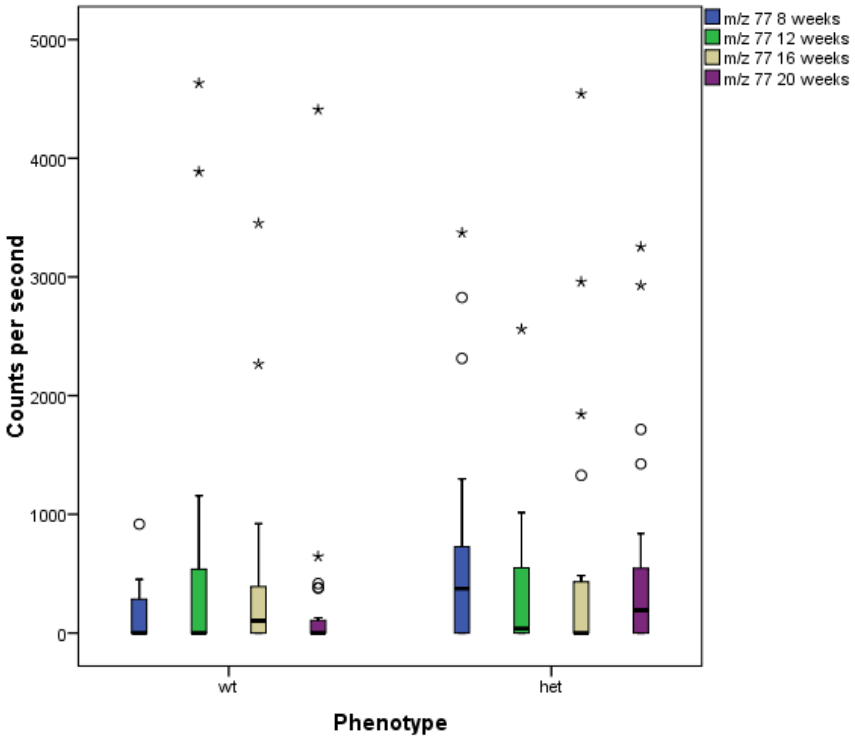


Figure 5. 17 – Boxplots of m/z 77; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

Correlations and age effect in faecal headspace profile

Positive medium correlations were found between *m/z* counts and blood glucose levels:

Correlation for *m/z* 45 (ethanol) at 20 weeks of age (r = 0.301, n = 47, p = 0.040)

Correlation for *m/z* 47 (unknown) at 8 weeks of age (r = 0.338, n = 47, p = 0.020)

Correlation for *m/z* 48 ($NO^+(H_2O)$) at 20 weeks of age (r = 0.310, n = 47, p = 0.034)

A Friedman test (repeated measures) suggest that are significant differences in *m/z* 48 ($NO^+(H_2O)$) across the age ($p < 0.001$). Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a significant difference between 8 to 12 weeks

of age (z = -2.868, p = 0.004); 12 to 16 weeks of age (z = -2.836, p = 0.005); and 8 to 20 weeks of age (z = -2.0656, p = 0.08).

SIFT-MS O_2^+ data set: the Mann-Whitney U test revealed statistically significant differences in the VOCs levels of mutant (het) mice and WT littermates, these were identified and are presented in Table 5.3.

Table 5. 3 – Median values for statistically significant m/z values tested over the age (8, 12, 16 and 20 weeks old) of faecal headspace in Cushing´s mice and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test (p < 0.05).

			Median		
	Ion	***Possible compound (s)***	***wt n = 35***	***het n = 30***	*Significance level (p)*
8 weeks	*m/z* 43	Acetone, acetaldehyde, 2-butanone, 2-pentanone, acetic acid	1.5×10^3	2.4×10^3	0.023
	m/z 60	Unknown	1.8×10^2	6.4×10^2	0.006
	m/z 74	Unknown	0.0	2.3×10^2	0.002
	m/z 78	Benzene	0.0	4.6×10^2	0.008
	m/z 92	Toluene	0.0	2.2×10^2	0.019
	Ion	***Possible compound (s)***	***wt n = 23***	***het n = 24***	***Significance level (p)***
12 weeks	*m/z* 43	Acetone, acetaldehyde, 2-butanone, 2-pentanone, acetic acid	2.4×10^3	4.4×10^3	0.010
	m/z 60	Unknown	7.1×10^2	3.0×10^3	0.005
	m/z 61	Unknown	4.4×10^1	3.1×10^2	0.012
	m/z 78	Benzene	0.00	7.4×10^2	0.032
	Ion	***Possible compound (s)***	***wt n = 34***	***het n = 30***	***Significance level (p)***
16 weeks	*m/z* 105	Benzaldehyde	0.0	0.0	0.030
	Ion	***Possible compound (s)***	***wt n = 35***	***het n = 30***	***Significance level (p)***
20 weeks	*m/z* 31	Isopropylamine	4.1×10^3	7.2×10^3	0.040
	m/z 74	Unknown	0.0	1.7×10^2	0.037
	m/z 92	Toluene	87.60	3.9×10^2	0.044

SIFT-MS O_2^+ data set confirms the presence of significant greater levels of acetone (*m/z* 43) and acetic acid (*m/z* 43) in faecal headspace of mutant (het) mice. A graphical representation is given below.

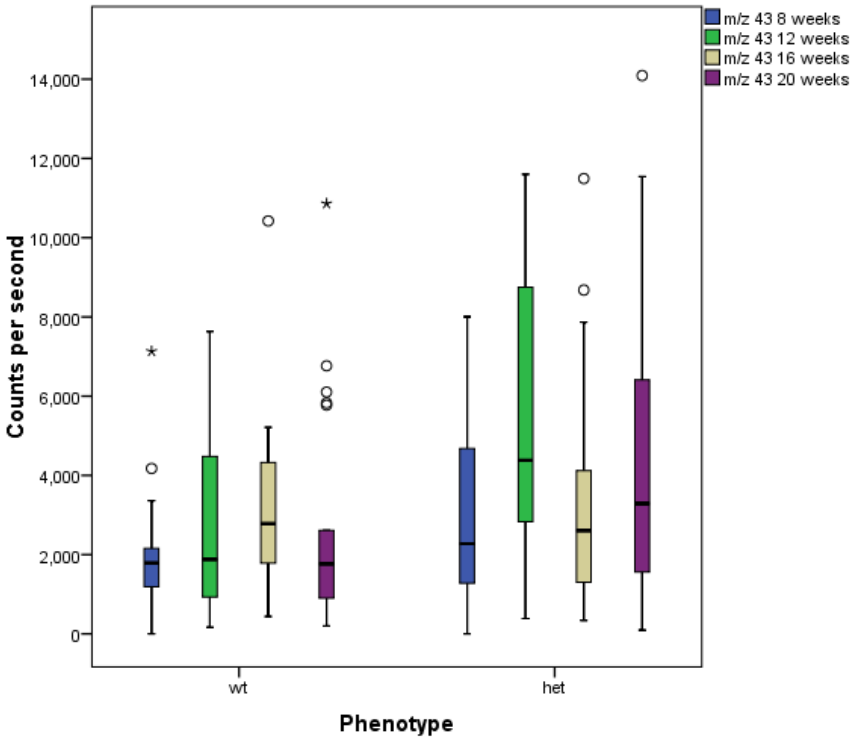


Figure 5. 18 – Boxplots of m/z 43; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

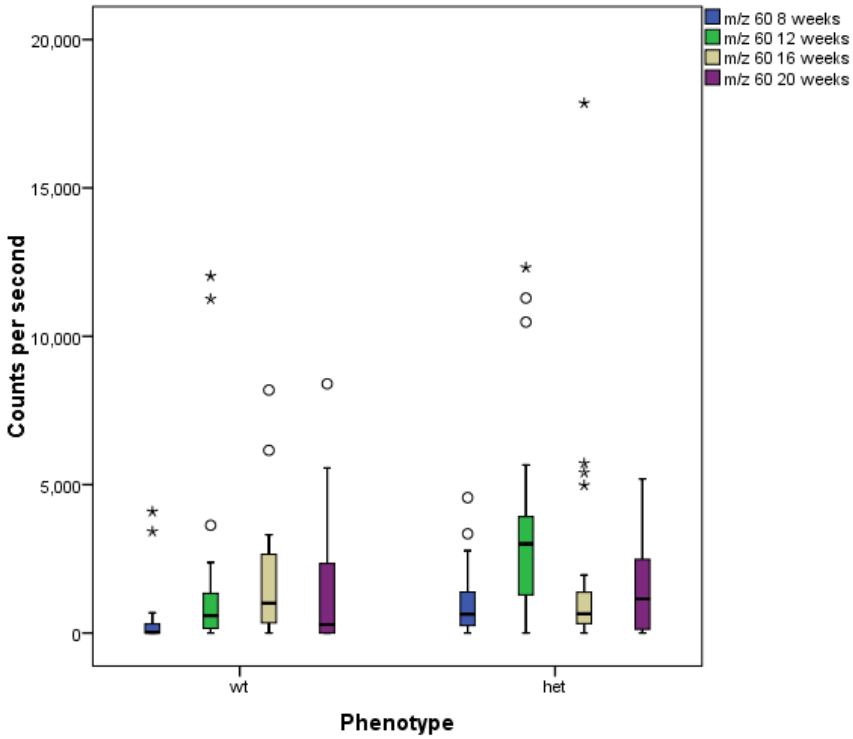


Figure 5. 19 – Boxplots of m/z 60; Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

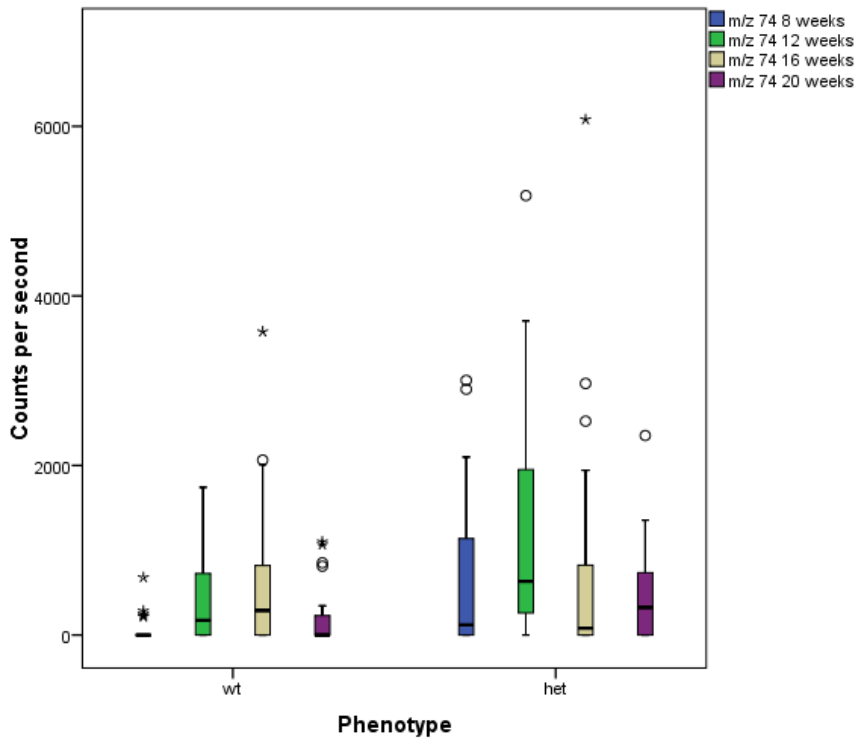


Figure 5. 20 – Boxplots of m/z 74; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

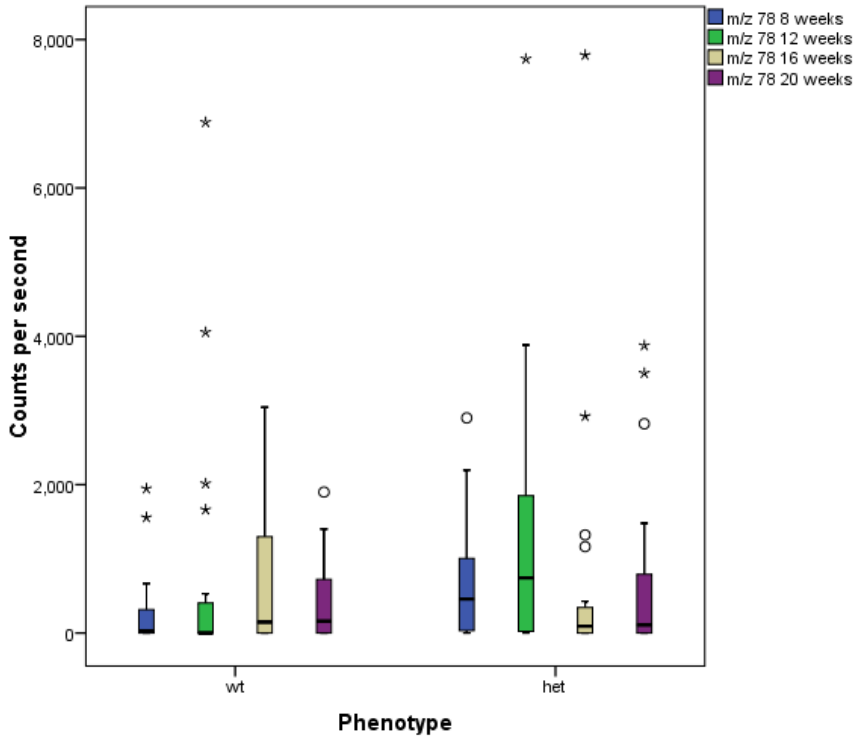


Figure 5. 21 – Boxplots of m/z 78; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

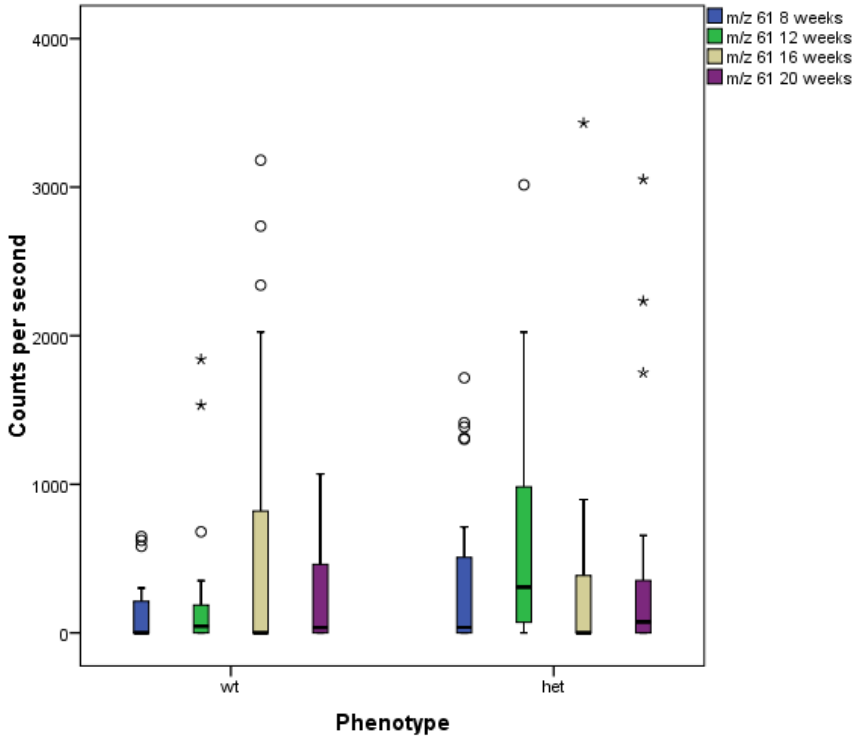


Figure 5. 22 – Boxplots of m/z 61; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

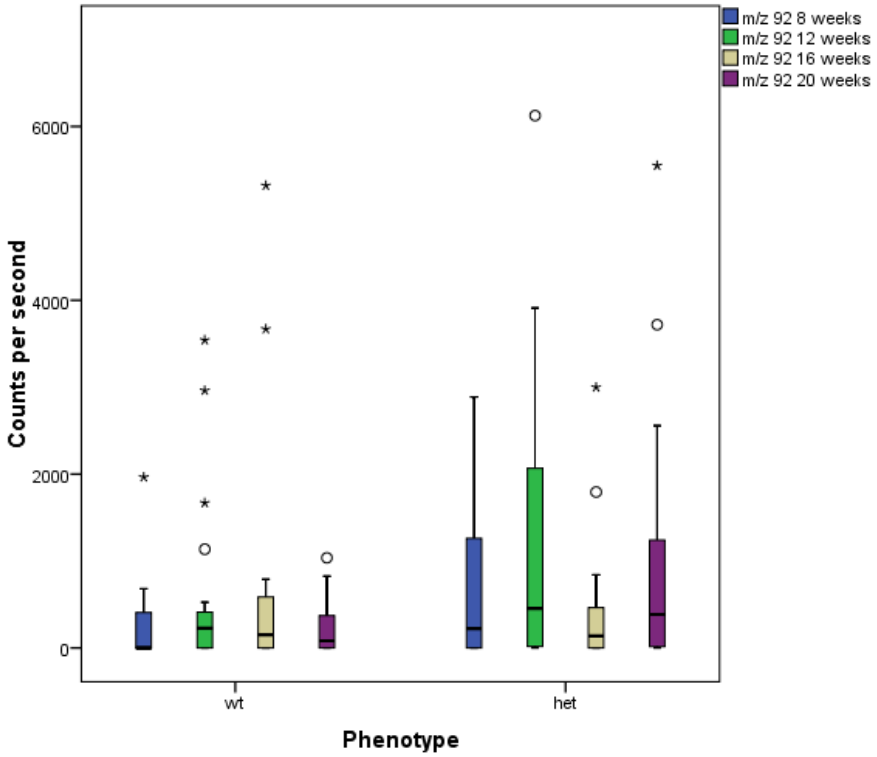


Figure 5. 23 – Boxplots of m/z 92; Faecal headspace of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

A weak but positive correlation (r = 0.303, n = 46, p = 0.041) was found between *m/z* 60 (which is currently unassigned compound) and blood glucose levels at 20 weeks of age, using the non-parametric Spearman's rho correlation coefficient. In addition to a significant change of VOC over the age (p < 0.001), where Wilcoxon Signed Rank Test using a Bonferroni adjusted p value revealed a significant difference between 8 to 12 weeks of age (z = -4.540, p = 0.000).

5.1.2 Multivariate Statistics using SIFT-MS data

A discriminant analysis was conducted for each data set individually using a single precursor ion at different time points (8, 12, 16 and 20 weeks of age); by combining all data sets at different time points for each precursor ion; and by combining data sets using H_3O^+, NO^+ and O_2^+ to predict whether an animal was within the wild-type (WT) group or mutant (het) group. The cross-validated classification is presented in Table 5.4 and the corresponding test sensitivity and specificity was determined.

Table 5. 4 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity given for faecal headspace of Cushing´s mice data set using SIFT-MS. The best classification results are highlighted in bold.

Precursor ion	*Age (weeks)*	*Cross-validated classification (%)*	*Sensitivity (%)*	*Specificity (%)*
H_3O^+	**8**	**89.2**	**87**	**91**
	12	81.5	77	86
	16	75.4	67	83
	20	**86.2**	**83**	**89**
	all	71.9	68	76
NO^+	**8**	**83.1**	**77**	**89**
	12	**80.9**	**79**	**83**
	16	73.8	50	94
	20	70.8	70	71
	all	75.2	65	84
O_2^+	8	81.5	73	89
	12	**85.1**	**83**	**87**
	16	**92.2**	**90**	**94**
	20	72.3	57	86
	all	73.0	68	78
$H_3O^+ + NO^+ + O_2^+$	all	65.9	58	73

Discriminant analysis was carried out to predict whether an animal was within the mutant (het) group or WT group. The best classification results are highlighted in bold in Table 5.4. The cross-validated classification showed that overall 89.2% were correctly classified using H_3O^+ data set at 8 weeks old and 86.2% were correctly classified using H_3O^+ data set at 20 weeks. Sensitivities and specificities are given respectively. The analytical method used showed a 83.1% correct classification using NO^+ data set at 8 weeks old and 80.9% were correctly classified using NO^+ data set at 12 weeks old. The best classification results are given for using O_2^+ data set at 12 and 16 weeks old, where the cross-validated classification showed an overall 85.1% and 92.2% correct classification respectively. Combining data sets ($H_3O^+ + NO^+ + O_2^+$) to discriminate between the groups (wt/het) showed that overall 65.9% were correctly classified, with a sensitivity of 58% and specificity of 73%.

According to the previous findings, the best set of predictors that discriminate between the groups are indicated in the following Tables 5.5, 5.6 and 5.7. Wilks´ Lambda indicates the significance of the discriminant function, thus all predictors are significant ($p < 0.000$).

Table 5. 5 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS H_3O^+ data set.

Precursor ion/ Age	*Variables (m/z)*	*Possible compound (s)*
H_3O^+ (8 weeks)	54	ammonia
	57	Propanoic acid, butanol
	71	Butanoic acid, pentanol
	79	Acetic acid, propanol
	122	Unknown
H_3O^+ (20 weeks)	38	Unknown
	60	Isopropylamine, methylethylamine, 1-propylamine
	63	Acetaldehyde
	65	Ethanol
	97	Acetic acid, propanol
	100	Unknown

Table 5. 6 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS NO^+ data set.

Precursor ion/ Age	*Variables*	*Possible*
	m/z	*compound (s)*
NO^+ (8 weeks)	32	Artefact of O_2^+
	43	Acetaldehyde
	62	Dimethyl sulphide, methanol
	104	Propanoic acid
	118	Butanoic acid
NO^+ (12 weeks)	47	Unknown
	66	Unknown
	69	Unknown
	89	Unknown
	104	Propanoic acid

Table 5. 7 – Best set of predictors discriminating between WT littermates and mutant (het) Cushing´s mice using SIFT-MS O_2^+ data set, and combining SIFT-MS $H_3O^+ + NO^+ + O_2^+$.

Precursor ion/ Age	*Variables*	*Possible*
	m/z	*compound (s)*
O_2^+ (12 weeks)	43	Acetic acid, acetone, 2-butanone, 2-pentanone, acetaldehyde
	46	Formic acid
	60	Unknown
	72	2-butanone, butanal, hexanal
	104	Unknown
O_2^+ (16 weeks)	69	Pentanal
	77	Unknown
	81	Unknown
	108	Methyl phenol
	120	Unknown

5.1.3 Univariate Analysis using GC-MS data

Quantification using GC-MS was performed using an internal standard d8-toluene loaded onto each sample tube. Although quantification using this method is only semi-quantitative, GC-MS results might validate and support the SIFT-MS findings discussed above.

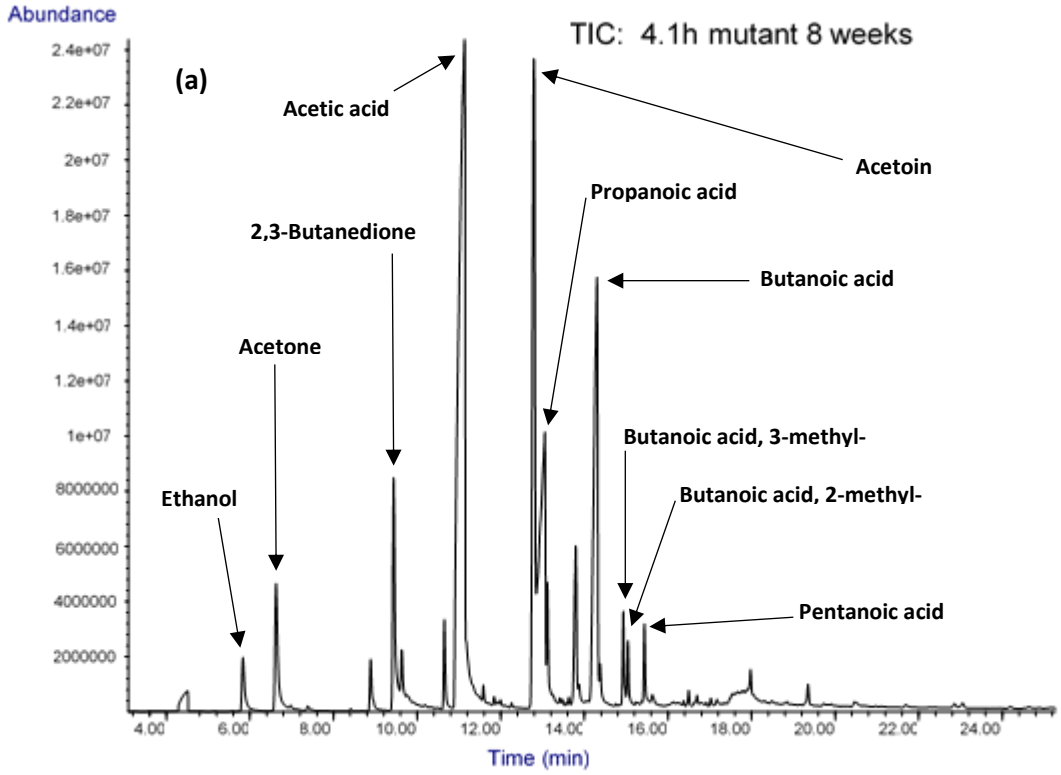
A cohort of 65 mice (21 males and 44 females) was used of whom 30 were genotyped as mutants and 35 WT littermates. Deconvolution and identification of the top ten compounds present in the faecal headspace of Cushing´s mice is given in Table 5.8 and

the retention time in the column respectively. Highlighted in bold are the three most abundant VOCs in faecal headspace of Cushing´s mice at each time period.

Table 5. 8 – Quantification of top 10 abundant compounds in faecal headspace of Cushing´s mice (8, 12, 16 and 20 weeks of age). Data acquired by GC-MS using internal standard addition of d8-toluene. Retention time (RT) is given in minutes. Highlighted in bold is given the three most abundant VOCs in faecal headspace of Cushing´s mice.

CAS	***RT [min]***	***Compound***	***Mean concentration 8 weeks ($ng\ L^{-1}$)***	***Mean concentration 12 weeks ($ng\ L^{-1}$)***	***Mean concentration 16 weeks ($ng\ L^{-1}$)***	***Mean concentration 20 weeks ($ng\ L^{-1}$)***
64197	11.13	**Acetic acid**	**6.8×10^2**	**7.0×10^2**	**9.3×10^2**	**8.8×10^2**
107926	14.30	**Butanoic acid**	**3.2×10^2**	**2.5×10^2**	**4.4×10^2**	**4.0×10^2**
513860	12.80	**Acetoin**	**3.0×10^2**	**2.1×10^2**	**3.7×10^2**	**3.5×10^2**
79312	14.54	Propanoic acid, 2-methyl-	1.9×10^2	1.6×10^2	1.8×10^2	3.0×10^2
79094	13.05	Propanoic acid	2.0×10^2	**2.0×10^2**	3.0×10^2	2.4×10^2
503742	14.92	Butanoic acid, 3-methyl-	9.0×10^1	1.0×10^2	1.1×10^2	1.1×10^2
116530	15.04	Butanoic acid, 2-methyl-	8.6×10^1	1.1×10^2	1.2×10^2	1.1×10^2
431038	9.43	2,3-butanedione	8.8×10^1	7.2×10^1	1.1×10^2	1.1×10^2
71363	11.58	1-Butanol	7.7×10^1	9.4×10^1	1.4×10^2	1.0×10^2
109524	15.43	Pentanoic acid	9.0×10^1	8.3×10^1	1.4×10^2	1.0×10^2

Overall, acetic acid, butanoic acid and acetoin are present in faecal headspace in greater concentration than other compounds (Table 5.8). In Figure 5.24 present an illustrative chromatogram of a mutant animal at 8 weeks of age **(a)** and showing the increase in VOC across the age (8 and 20 weeks) **(b)**.



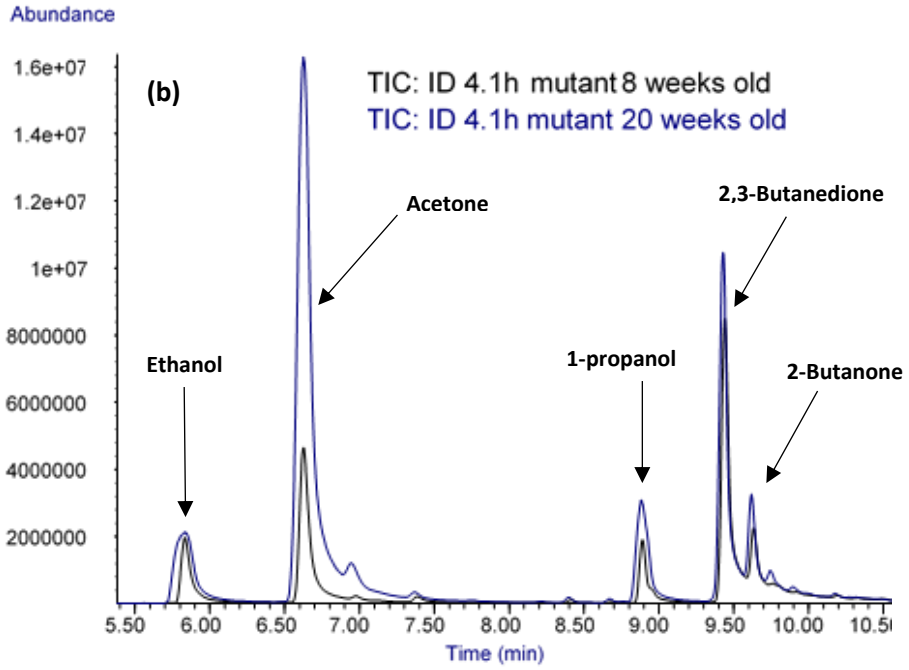


*Figure 5. 24 – Illustrative chromatogram of a mutant (het) animal at 8 weeks of age **(a)**; zoom in overlapping chromatograms at 8 weeks and 20 weeks of age highlighted in blue, the same animal ID showing the increase in VOC across the age **(b)**.*

Table 5. 9 – Median concentrations expressed in ng L^{-1} for statistically significant VOCs tested over the age (8, 12, 16 and 20 weeks of age) of Cushing´s mice. Highlighted in bold is given the compounds that are statistically significant for all time points. Data acquired by GC-MS and analysed using Mann Whitney´s test (p < 0.05).

Age (weeks)	***CAS***	***RT [min]***	***Compound***	***Median concentration***	
				wt n = 35	***het n = 30***
8	64197	11.13	**Acetic acid**	**3.0×10^2**	**8.7×10^2**
	431038	9.43	2,3-butanedione	8.2×10^1	6.5×10^1
	79094	13.05	**Propanoic acid**	**8.9×10^1**	**2.2×10^2**
	71363	11.58	1-butanol	3.0×10^1	4.7×10^1
	116530	15.04	Butanoic acid, 2-methyl-	3.3×10^1	5.8×10^1
12	64197	11.13	**Acetic acid**	**3.7×10^2**	**8.4×10^2**
	79094	13.05	**Propanoic acid**	**1.2×10^2**	**1.7×10^2**
	79312	14.54	Propanoic acid, 2-methyl-	5.5×10^1	7.6×10^1
	71363	11.58	1-butanol	3.3×10^1	7.2×10^1
	109524	15.43	Pentanoic acid	3.9×10^1	5.9×10^1
	503742	14.92	Butanoic acid, 3-methyl-	5.2×10^1	7.3×10^1
16	64197	11.13	**Acetic acid**	**6.3×10^2**	**1.0×10^3**
	431038	9.43	2,3-butanedione	9.6×10^1	5.7×10^2
	79094	13.05	**Propanoic acid**	**1.6×10^2**	**2.7×10^2**
	79312	14.54	Propanoic acid, 2-methyl-	5.8×10^1	1.1×10^2
	71363	11.58	1-butanol	6.0×10^1	1.3×10^2
20	64197	11.13	**Acetic acid**	**5.3×10^2**	**1.1×10^3**
	79094	13.05	**Propanoic acid**	**1.3×10^2**	**2.5×10^2**
	109524	15.43	Pentanoic acid	6.7×10^1	1.2×10^2
	116530	15.04	Butanoic acid, 2-methyl-	6.8×10^1	1.0×10^2

The chromatographic data do not follow a normal distribution, therefore the non-parametric Mann Whitney U test was used at each time point and revealed statistically significant differences ($p < 0.05$) in VOC concentration in faecal headspace of mutant (het) Cushing´s mice and WT littermates and these are indicated in Table 5.9. Remarkably, differences in acetic acid and propanoic acid concentration are statistically significant at each time point, and these occur in significantly increased levels in mutant mice, which is in agreement with previous findings discussed earlier using univariate analysis and multivariate statistics via SIFT-MS (Figure 5.25). However, the results of the Friedman test (repeated measures) suggest that are no significant differences in acetic

acid and propanoic acid levels across the age range nor are there any significant differences in the VOC profile concentration between males and females.

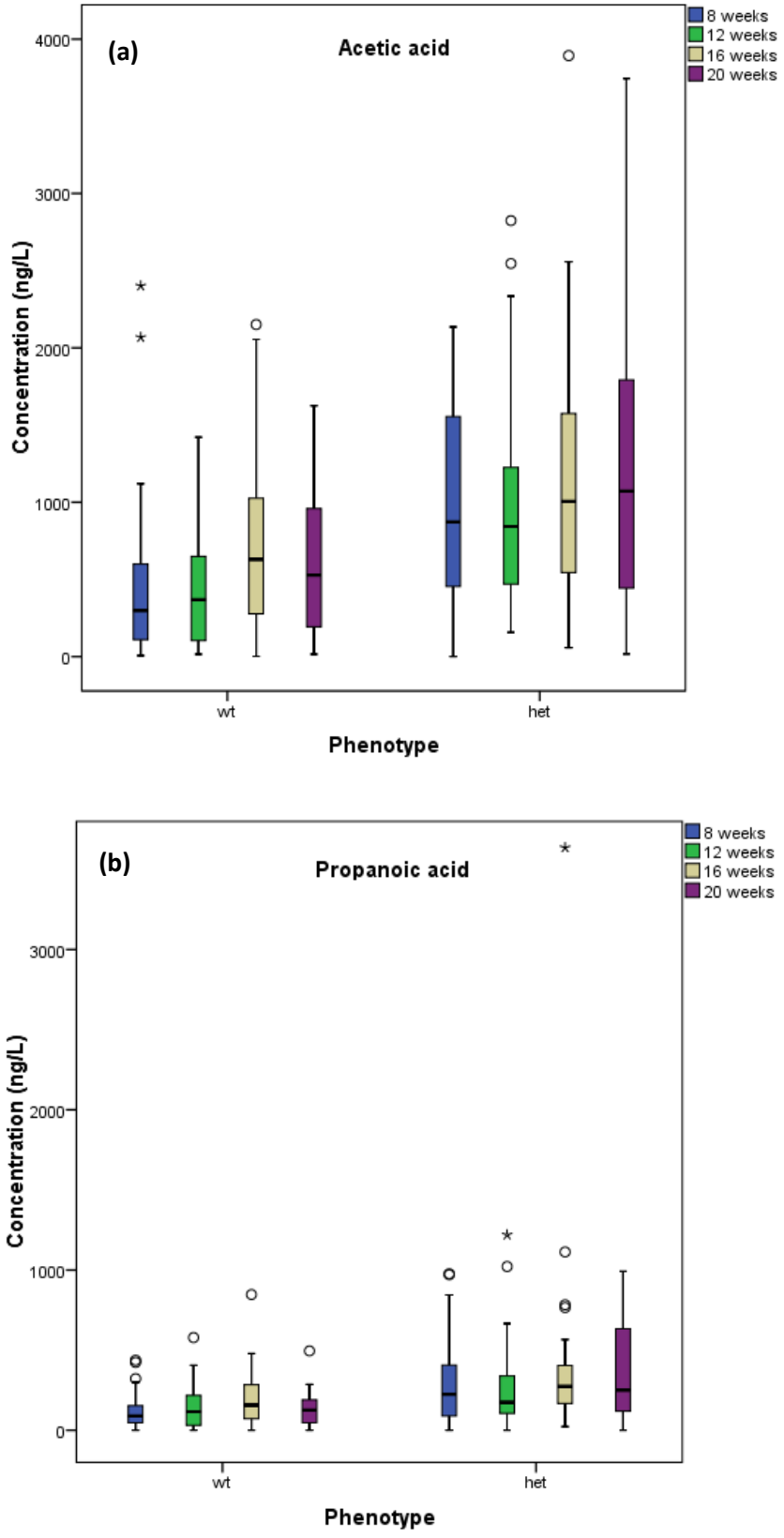


*Figure 5. 25 – Acetic acid **(a)** and propanoic acid **(b)** faecal headspace concentration of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data acquired by GC-MS and analysed using Mann Whitney's test ($p < 0.05$).*

Figure 5.26 present a graphical representation of statistically significant 2,3-butanedione **(a)**, propanoic acid, 2-methyl- **(b)**, 1-butanol **(c)**, pentanoic acid **(d)**, butanoic acid, 3-methyl- **(e)**, and butanoic acid, 2-methyl- **(f)** of Cushing´s mice (het) compared with WT littermates on a B6 C3PDE background. Mutant animals had significantly increased concentrations of butanoic acid, 3-methyl- at 12 weeks of age. Significantly greater concentrations of 1-butanol were found in mutant animals at 8, 12 and 16 weeks of age. Mutant animals had significantly increased concentrations of butanoic acid, 2-methyl- at 8 and 20 weeks of age; pentanoic acid at 12 and 20 weeks of age; and propanoic acid, 2-methyl- at 12 and 16 weeks of age. Interestingly, the volatile 2,3-butanedione was found to exist in significantly higher levels in WT littermates than mutant mice at 8 and 16 weeks of age.

Although not statistically significant at any time point, a Friedman test (repeated measures) suggests that are differences in acetoin levels across the age ($p < 0.001$). A Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a statistically significant fluctuation of acetoin with age increase, i.e. from 8 to 12 weeks *($z = -3.73, p < 0.001$);* and from 12 to 16 weeks *($z = -4.60, p < 0.001$).* The last time point (16 to 20 weeks) is not significant.

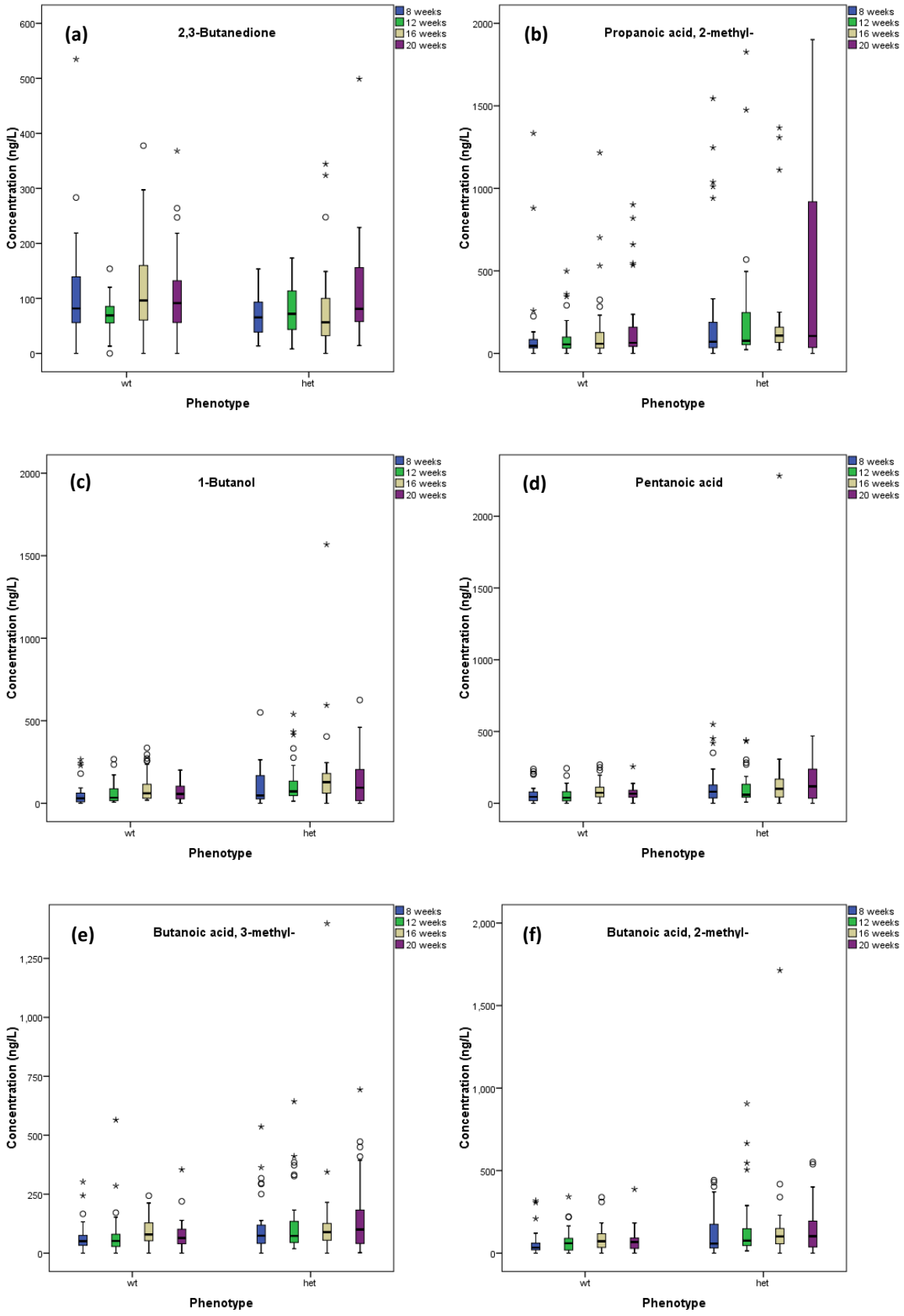


*Figure 5. 26 – Faecal headspace concentration of 2,3-butanedione **(a)**, propanoic acid, 2-methyl- **(b)**, 1-butanol **(c)**, pentanoic acid **(d)**, butanoic acid, 3-methyl- **(e)**, and butanoic acid, 2-methyl- **(f)** of Cushing´s mice (het) compared with WT littermates on a B6-C3PDE background along the age. Data analysed using Mann Whitney's test ($p < 0.05$).*

The relationship between VOC concentration and blood glucose levels was investigated using the non-parametric Spearman's rho correlation coefficient. Outliers have not been removed throughout this study. A small correlation *(r = 0.236, n = 65, p = 0.059)* was found between acetic acid concentration and blood glucose levels at 20 weeks of age.

5. 2 A longitudinal study of the VOC profile emitted by the faecal headspace of single Afmid knockout mice exhibiting impaired glucose tolerance

As discussed earlier in chapter 1, *Afmid* mice are prone to develop impaired glucose intolerance (IGT) which is a pre-diabetic state of hyperglycaemia. People with IGT have blood glucose levels that are higher than normal but not high enough to say they have diabetes. As discussed in chapter 1, people with IGT have plasma levels that are between 7.8 to 11.0 mmol L^{-1} glucose.

The disease was tracked across the three time points (8, 12 and 16 weeks of age) for the mutant (hom) *Afmid* mice and the WT littermates. The body weight of the animals is not normally distributed therefore a Mann Whitney U test was used to compare the mean score of weights. The mutation had no significant effects on body weight (Figure 5.27).

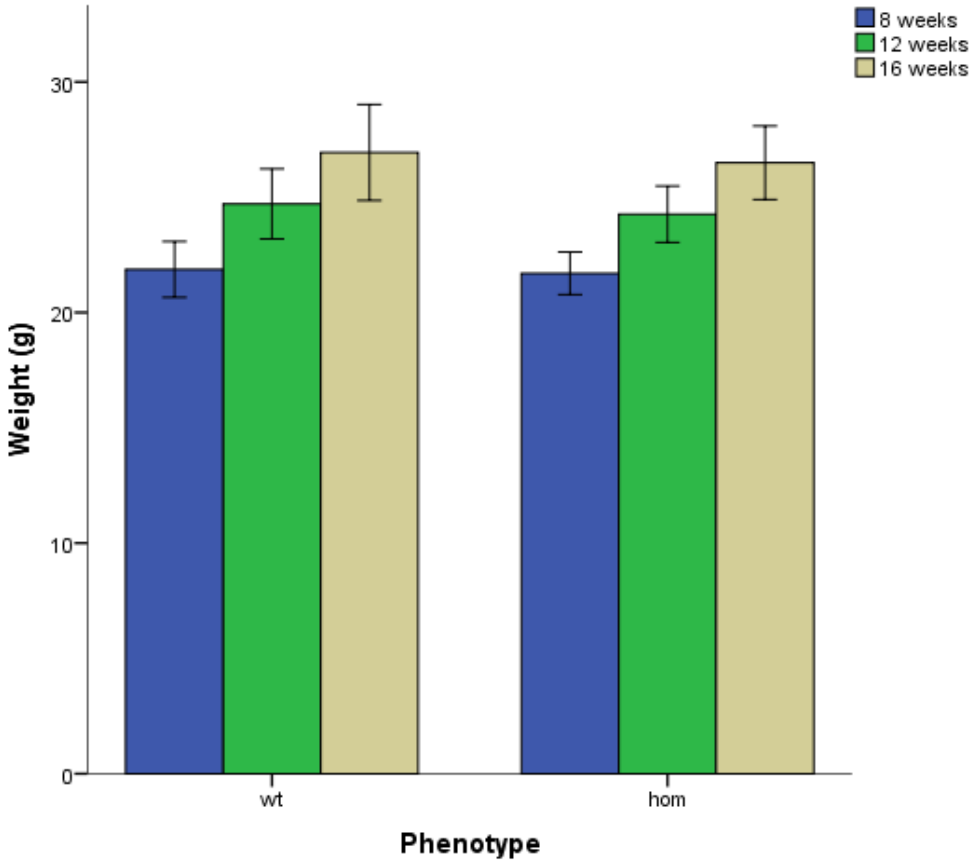


Figure 5. 27 – Body weights along the age of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test.

The non-parametric Friedman test (repeated measures) combined with individual post-hoc tests, i.e. Wilcoxon Signed Rank Test using a Bonferroni adjusted p value ($p < 0.01$), was used to compare the animal's weight along the age and to establish the statistically significant difference among the three time points. The results indicated that the body weight of the animals significantly change over the age ($p < 0.001$) and across the three time points, although there is no significant difference in the weight between mutants (hom*)* and WT littermates according to Mann Whitney's test (Figure 5.28). Males [*8 weeks (median = 2.4 × 10^1, $p < 0.05$); 12 weeks (median = 2.7 × 10^1, $p < 0.05$); 16 weeks (median = 2.9 × 10^1, $p < 0.05$)*] are significantly heavier than females [*8 weeks (median = 2.0 × 10^1, $p < 0.05$); 12 weeks (median = 2.2 × 10^1, $p < 0.05$); 16 weeks (median = 2.3 × 10^1, $p < 0.05$)*] over the three periods of time.

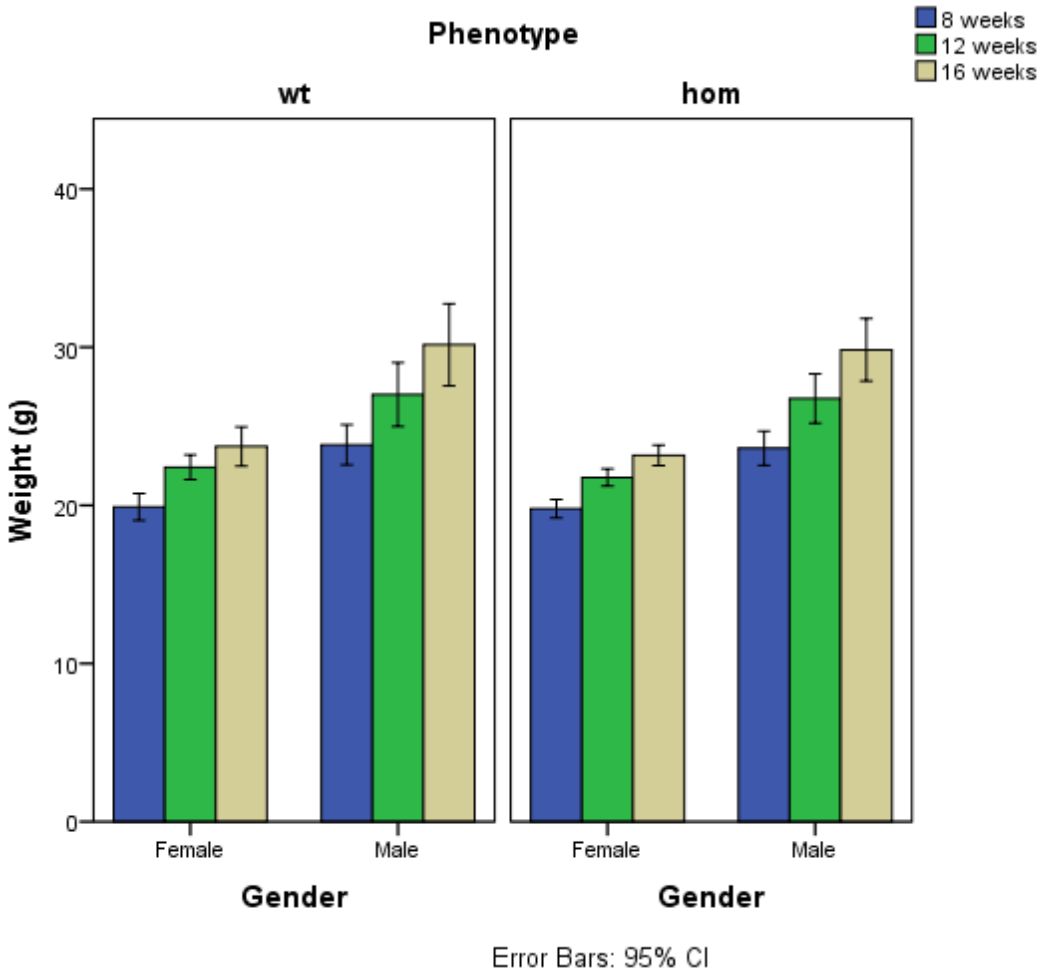


Figure 5. 28 – Body weights along the age of male and female Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using non-parametric Friedman test ($p < 0.001$) together with Wilcoxon Signed Rank Test using a Bonferroni adjusted p value ($p < 0.01$).

The variable blood glucose is not normally distributed therefore, individual Mann Whitney U tests were conducted to compare the mean scores per each time point. There is a significant difference in blood glucose levels of homozygous (hom) mice and WT littermates at 8, 12 and 16 weeks of age. Mutant (hom) mice developed significant increased concentrations of blood glucose (Figure 5.29).

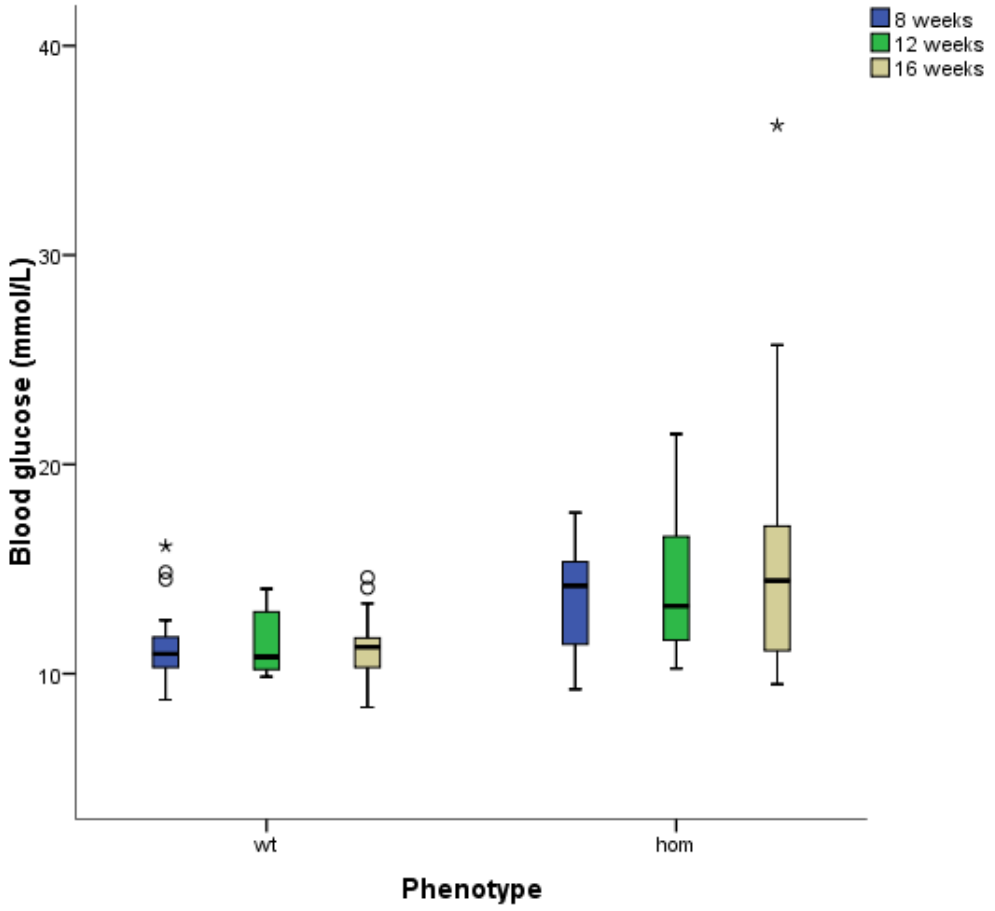


Figure 5. 29 – Blood glucose concentrations along the age in Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$).

Blood glucose was taken for unfasted animals with the blood being sampled at the same period of the day, between 2 pm and 4 pm.

The mutant (hom) male mice developed significant increased concentrations of blood glucose ($p < 0.05$) at the three time points compared to the WT littermates (Figure 5.30), although the increase of blood glucose levels with age are not statistically significant.

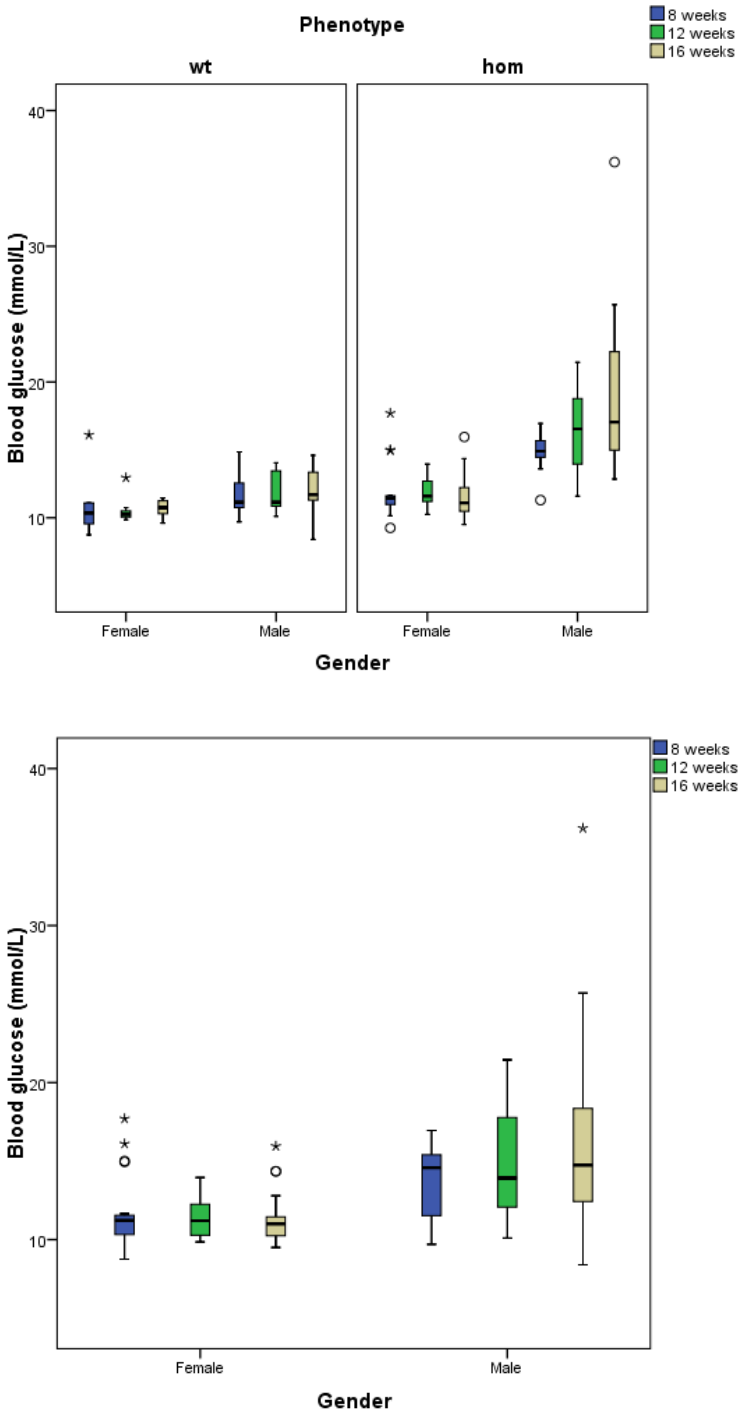


Figure 5. 30 – Blood glucose concentrations along the age of male and female Afmid (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$).

The plasma insulin concentrations indicated in Figure 5.31 do not correspond to 16 weeks of age but a later stage once the animals were culled, i.e. animals born in March were culled at 32 weeks of age, and animals born in May were culled at 28 weeks of age. Mann Whitney U test revealed no significant difference in plasma insulin concentration of

mutant (hom) mice and WT littermates (Figure 5.31) which is in agreement with previous findings (Hugill *et al.*, 2015). As discussed earlier in chapter 1, these mice show impaired glucose tolerance, although their insulin sensitivity is unchanged when compared to wild-type animals (Hugill *et al.*, 2015).

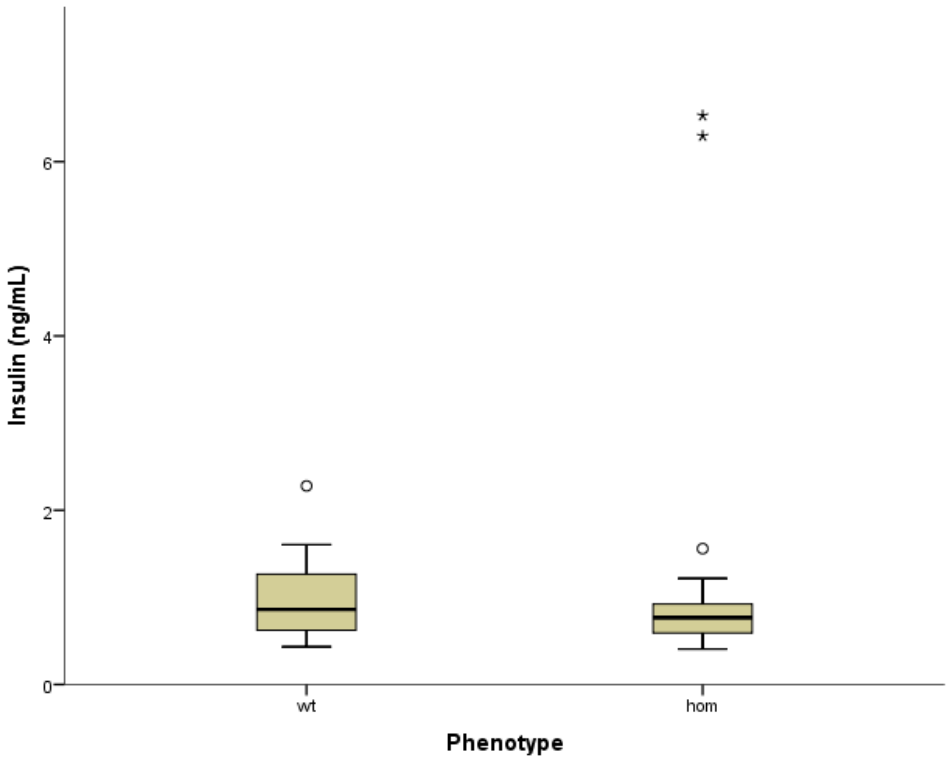


Figure 5. 31 – Plasma insulin concentrations of Afmid (hom) compared with WT littermates on a C57BL6/NTac background.

5.2.1 Univariate Analysis using SIFT-MS data

The faecal volatile profile of *Afmid* mice was acquired using SIFT-MS at The Open University and the univariate analysis is given in this chapter.

The SIFT-MS data set do not follow normal distribution either for H_3O^+, NO^+ and O_2^+ precursor ions. According to the previous findings, a non-parametric statistic technique, Mann-Whitney U Test, was applied to test the differences between the two independent

groups of variables, i.e. mutant (hom) mice and WT littermates, and the same test was applied at each time point. The Mann-Whitney U test revealed significant differences in the VOCs levels of mutants (hom) animals and WT littermates, either for H_3O^+, NO^+ or O_2^+ precursor ions.

SIFT-MS H_3O^+ data set was obtained using a cohort of 48 mice (24 males and 24 females) of whom 30 were genotyped as mutants and 18 WT littermates. Using H_3O^+ as the precursor ion, statistically significant differences in the VOCs levels of mutant (hom) mice and WT littermates, were identified and given in Table 5.10.

Below is shown an illustrative mass spectrum (Figure 5.32) zoomed in of one sample from a mutant (hom) mouse. Data acquired by SIFT-MS H_3O^+, *m/z* range 10-140, acquisition time of 5 seconds and recorded for six iterations.

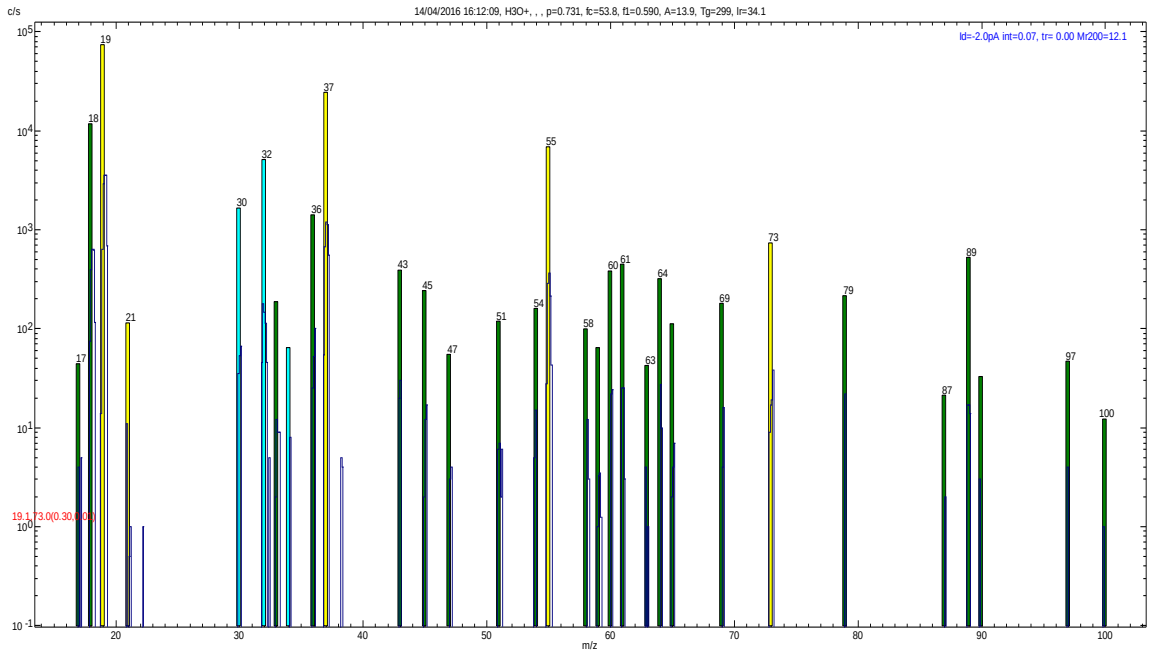


Figure 5. 32 – Illustrative mass spectrum of a faecal headspace sample from a mutant (hom) Afmid mouse (animal ID 51.1c). Data acquired by SIFT-MS H_3O^+, m/z range 10-140, acquisition time of 5 seconds and six iterations.

Table 5. 10 – Median values for statistically significant m/z values tested over the age (8 and 16 weeks of age) of faecal headspace Afmid mice and its significance levels acquired using H_3O^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$).

			Median		
		Possible compound (s)	**wt**	**hom**	
	Ion		***n = 18***	***n = 30***	*Significance level (p)*
8 weeks	*m/z* 59	Acetone, propanal	7.1×10^2	1.1×10^2	0.029
16 weeks	*m/z* 60	Acetonitrile	0.0	1.1×10^2	0.043

The ion *m/z* 59 is likely to be acetone and probably arise from-fatty acid and carbohydrate metabolism. The ion *m/z* 60 might not be an endogenous VOC. Boxplots are indicated in Figure 5.33.

No significant differences were found for 12 weeks of age.

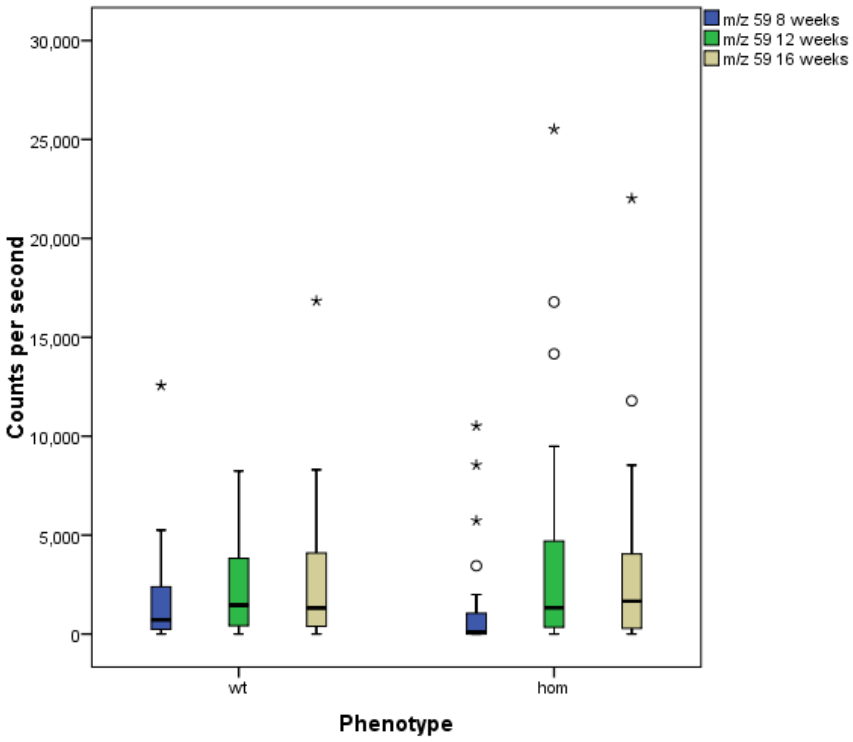


Figure 5. 33 – Boxplots of m/z 59; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).

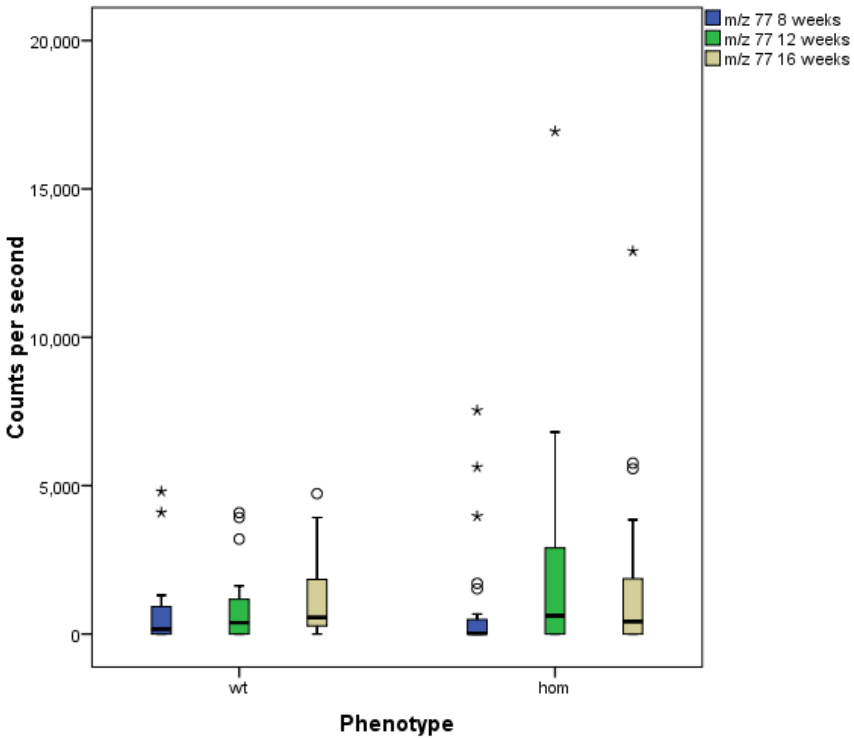


Figure 5. 34 – Boxplots of m/z 77; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).

Correlations and age effect in faecal headspace profile

The non-parametric Friedman test (repeated measures) followed with individual post-hoc tests, i.e. Wilcoxon Signed Rank Test using a Bonferroni adjusted p value ($p < 0.01$), was conducted and the results indicated that the ion *m/z* 59 (most likely acetone) significantly change over the age ($p < 0.001$) and specifically between 8 to 12 weeks of age. Additionally, a two medium correlations ($r = 0.304$, $n = 48$, $p = 0.036$) and ($r = 0.295$, $n = 48$, $p = 0.042$) were found between *m/z* 59 at 16 weeks and corresponding blood glucose levels, and *m/z* 77 at 16 weeks and corresponding blood glucose levels, respectively.

Although not statistically significant between mutant (hom) animals and WT littermates, the ion *m/z* 61 (potentially acetic acid or propanol) significantly changes over the age, particularly between 8 to 12 weeks and 8 to 16 weeks of age (Figure 5.35).

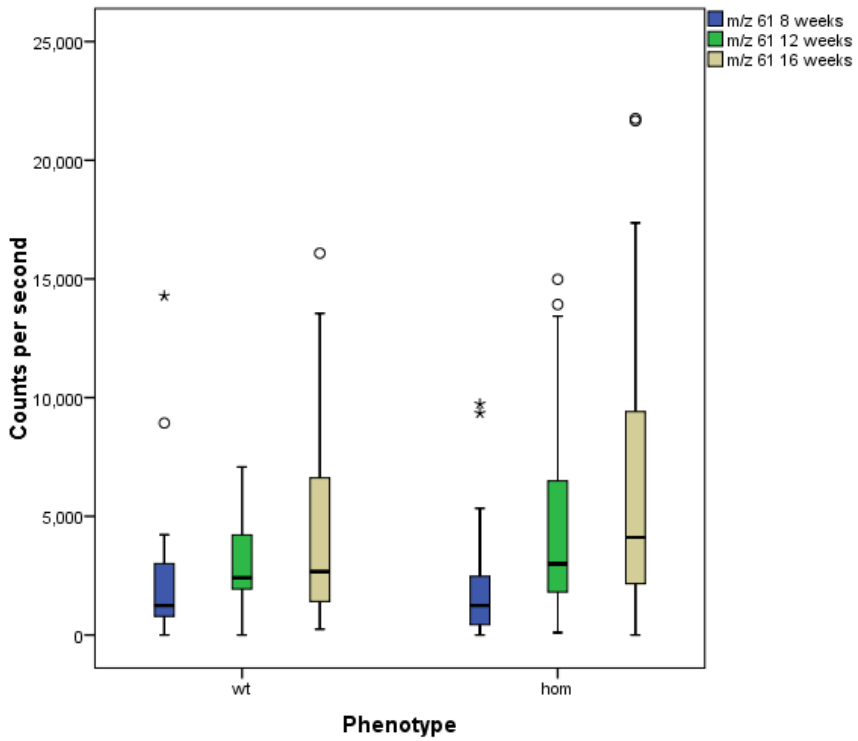


Figure 5. 35 – Boxplots of m/z 61; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS H_3O^+ and analysed using Mann Whitney's test ($p < 0.05$).

SIFT-MS NO^+ data set: the Mann-Whitney U test revealed statistically significant differences in the VOCs levels of mutant (hom) mice and WT littermates, were identified and presented in Table 5.11.

Table 5. 11 – Median values for statistically significant m/z values tested over the age (8, 12 and 16 weeks of age) of faecal headspace of Afmid mice and its significance levels acquired using NO^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$).

			Median		
8 weeks	*Ion*	***Possible compound (s)***	***wt*** ***n = 18***	***hom*** ***n = 30***	*Significance level (p)*
	m/z 66	Unknown	0.0	3.3×10^2	0.026
12 weeks	*m/z* 66	Unknown	0.0	4.5×10^2	0.013
	m/z 118	Butanoic acid	9.9×10^2	4.5×10^2	0.059
16 weeks	*m/z* 43	Acetaldehyde	0.0	3.4×10^1	0.054

The unknown ion *m/z* 66 was found in significantly greater levels in mutant (hom) mice ($p < 0.05$), Figure 5.36. This is probably an artefact $NO^+(H_2O)_2$.

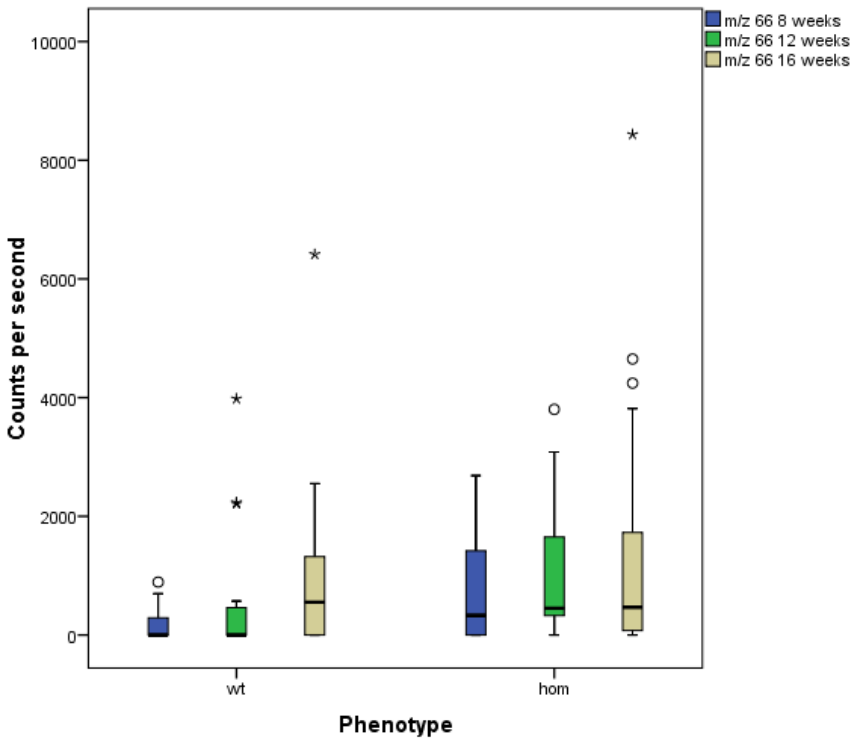


Figure 5. 36 – Boxplots of m/z 66; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background along the age. Data acquired by SIFT-MS NO^+ and analysed using Mann Whitney's test ($p < 0.05$).

SIFT-MS O_2^+ data set: Statistically significant differences were detected for *m/z* values indicated in Table 5.12, by means of using a Mann-Whitney U test. Median values for each VOC within the mutant (hom) and WT groups are indicated and its significance levels.

Table 5. 12 – Median values for statistically significant m/z values tested for 8 weeks of age of faecal headspace of Afmid mice and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test ($p < 0.05$).

			Median		
		Possible compound (s)	***wt***	***hom***	
	Ion		***n = 18***	***n = 30***	*Significance level (p)*
8 weeks	---	---			---
	m/z 56	Hexanal, propenal	0.0	1.6×10^2	0.039
	m/z 89	Unknown	0.0	5.4×10^1	0.029

No significant differences were found at 12 and 16 weeks of age via SIFT-MS O_2^+. Significant differences were detected for *m/z* 56 at 8 weeks of age.

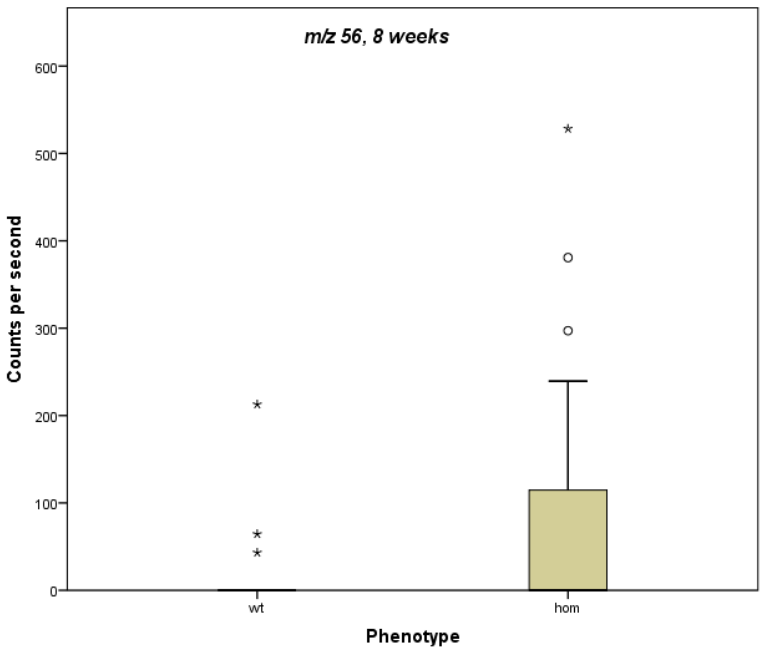


Figure 5. 37 – Boxplots of m/z 56; Faecal headspace of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background at 8 weeks of age. Data acquired by SIFT-MS O_2^+ and analysed using Mann Whitney's test ($p < 0.05$).

5.2.2 Multivariate Statistics using SIFT-MS data

A discriminant analysis was conducted for each *Afmid* data set individually using a single precursor ion at different time points (8, 12 and 16 weeks); by combining all data sets at different time points for each precursor ion; and by combining data sets using H_3O^+, NO^+ and O_2^+ to predict whether an animal was within the WT group or mutant (hom) group.

The cross-validated classification for all the precursor ions, and sensitivity and specificity are given in Table 5.13, where discriminant analysis was carried out to predict whether an animal was within the mutant (hom) group or WT group for each age point.

Table 5. 13 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity given for faecal headspace of Afmid mice data set using SIFT-MS. The best classification results are highlighted in bold.

Precursor ion	*Age (weeks)*	*Cross-validated classification (%)*	*Sensitivity (%)*	*Specificity (%)*
H_3O^+	**8**	**75.0**	**80**	**67**
	12	72.9	87	50
	16	77.1	97	44
	all	71.5	96	31
NO^+	8	77.1	83	67
	12	**95.8**	**97**	**94**
	16	77.1	97	45
	all	69.4	87	41
O_2^+	8	75.0	87	56
	12	72.9	87	50
	16	**91.7**	**93**	**89**
	all	78.5	91	57
$H_3O^+ + NO^+ + O_2^+$	all	68.3	92.6	27.8

The best classification results are highlighted in bold in Table 5.13. The cross-validated classification showed that overall 75.0% were correctly classified using H_3O^+ data set at 8 weeks of age. Sensitivities and specificities are given respectively. The analytical method used showed a 95.8% correct classification using NO^+ data set at 12 weeks of age. The best classification results are given for using O_2^+ data set at 16 weeks, where the cross-validated classification showed an overall 91.7% correct classification.

According to the previous findings, the best set of predictors that discriminate between the groups are indicated in the following Tables 5.14, 5.15 and 5.16. Wilks´ Lambda indicates the significance of the discriminant function, thus all predictors are significant ($p < 0.000$).

Table 5. 14 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS H_3O^+ data set.

Precursor ion/ Age	*Variables*	*Possible*
	m/z	*compound (s)*
H_3O^+ (8 weeks)	30	Unknown
	39	Unknown
	50	Unknown
	71	Butanoic acid, pentanol
	80	Pyridine

Table 5. 15 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS NO^+ data set.

Precursor ion/ Age	*Variables*	*Possible*
	m/z	*compound (s)*
NO^+ (12 weeks)	28	Unknown
	36	Unknown
	72	1-Butylamine
	78	Benzene
	134	Unknown

Table 5. 16 – Best set of predictors discriminating between WT littermates and mutant (hom) Afmid mice using SIFT-MS O_2^+ data set.

Precursor ion/ Age	*Variables*	*Possible*
	m/z	*compound (s)*
O_2^+ (16 weeks)	34	Unknown
	38	Unknown
	59	Methylethylamine
	79	Pyridine
	99	Unknown

5.2.3 Univariate Analysis using GC-MS data

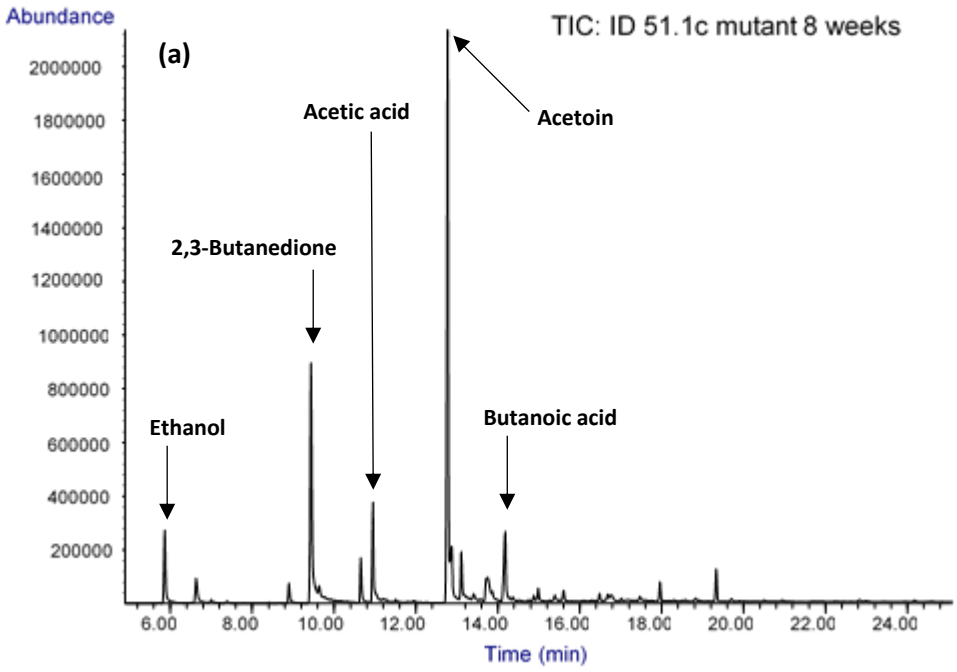
As the same manner, quantification using GC-MS was done by means of using an internal standard d8-toluene loaded onto each sample tube.

A cohort of 48 mice (24 males and 24 females) were bred of whom 30 were genotyped as mutants (hom) and 18 WT littermates. Four animals have been removed from the data set, thus, in this study was used a group of 44 mice (22 males and 22 females) whom 28 were genotyped as mutants (hom) and 16 WT littermates. Deconvolution and identification of the top ten compounds present in the faecal headspace of *Afmid* mice is given in Table 5.17. The three most abundant VOCs in faecal headspace of *Afmid* mice at each time period are highlighted in bold.

Table 5. 17 – Quantification of top 10 abundant compounds in faecal headspace of Afmid mice (8, 12 and 16 weeks of age). Data acquired by GC-MS using internal standard addition of d8-toluene. Retention time (RT) is given in minutes. Highlighted in bold is given the three most abundant VOCs in faecal headspace of Afmid mice.

CAS	***RT [min]***	***Compound***	***Mean concentration 8 weeks ($ng\ L^{-1}$)***	***Mean concentration 12 weeks ($ng\ L^{-1}$)***	***Mean concentration 16 weeks ($ng\ L^{-1}$)***
64197	11.13	**Acetic acid**	**1.8×10^3**	**5.7×10^2**	**1.2×10^3**
513860	12.80	**Acetoin**	**2.0×10^3**	**2.6×10^2**	**6.2×10^2**
107926	14.30	**Butanoic acid**	**1.1×10^3**	**2.6×10^2**	**5.8×10^2**
431038	9.43	2,3-butanedione	4.7×10^2	8.6×10^1	1.6×10^2
79094	13.05	Propanoic acid	2.3×10^2	1.4×10^2	2.6×10^2
79312	14.54	Propanoic acid, 2-methyl-	1.1×10^2	9.1×10^1	1.8×10^2
71363	11.58	1-Butanol	9.1×10^1	7.4×10^1	2.0×10^2
109524	15.43	Pentanoic acid	7.7×10^1	7.9×10^1	1.4×10^2
116530	15.04	Butanoic acid, 2-methyl-	8.8×10^1	8.4×10^1	1.5×10^2
503742	14.92	Butanoic acid, 3-methyl-	6.9×10^1	7.9×10^1	1.2×10^2

Overall, acetic acid, acetoin and butanoic acid are present in faecal headspace in greater concentration than other compounds. In Figure 5.38 is represented an illustrative chromatogram of a mutant (hom) mouse at 8 weeks of age **(a)** and showing the increase in VOC across the age (8 and 16 weeks) **(b)**.



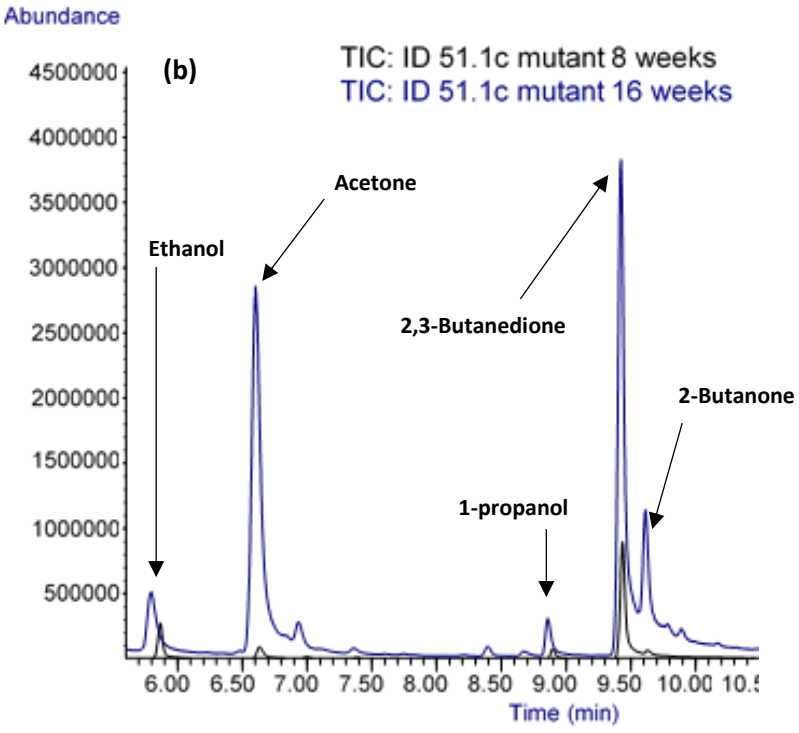


*Figure 5. 38 – Illustrative chromatogram of a faecal headspace sample of a mutant (hom) animal at 8 weeks of age **(a)**; zoom in overlapping chromatograms at 8 weeks and 16 weeks of age highlighted in blue, same animal ID showing the increase in VOC across the age **(b)**.*

The chromatographic data do not follow a normal distribution, therefore the non-parametric Mann Whitney U test was used at each time point (Table 5.18). Mann Whitney U test revealed no significant differences in VOC concentration in faecal headspace of mutant (hom) *Afmid* mice and WT littermates at 8 weeks of age. At 12 weeks of age, mutant (hom) animals had statistically significant elevated VOC concentrations in faecal headspace of acetic acid *(median = 6.2 × 10^2, n = 28),* propanoic acid *(median = 1.6 × 10^2, n = 28)* and 1-butanol *(median = 8.3 × 10^1, n = 28)* once compared to WT littermates respectively, acetic acid *(median = 4.1 × 10^2, n = 16),* propanoic acid *(median = 1.0 × 10^2, n = 16),* 1-butanol *(median = 6.6 × 10^1, n = 16), p < 0.05* (Figure 5.39). At 16 weeks of age, acetoin was found in significant increased levels in faecal headspace of mutant (hom) mice *(median = 5.4 × 10^2, n = 28),* WT littermates *(median = 4.0 × 10^2, n = 16), p < 0.05.*

There is no significant differences in the VOC concentration between males and females.

Table 5. 18 – Median concentrations expressed in ng L-1 for statistically significant VOCs tested over the age (8, 12 and 16 weeks of age) of faecal headspace of Afmid mice. Data acquired by GC-MS and analysed using Mann Whitney´s test (p < 0.05).

Age (weeks)	***CAS***	***RT [min]***	***Compound***	***Median concentration***	
				wt n = 16	***hom n = 28***
8	--	--	none	**--**	**--**
12	64197	11.13	Acetic acid	4.1×10^2	6.2×10^2
	79094	13.05	Propanoic acid	1.0×10^2	1.6×10^2
	71363	11.58	1-butanol	6.6×10^1	8.3×10^1
16	513860	12.80	Acetoin	4.0×10^2	5.4×10^2

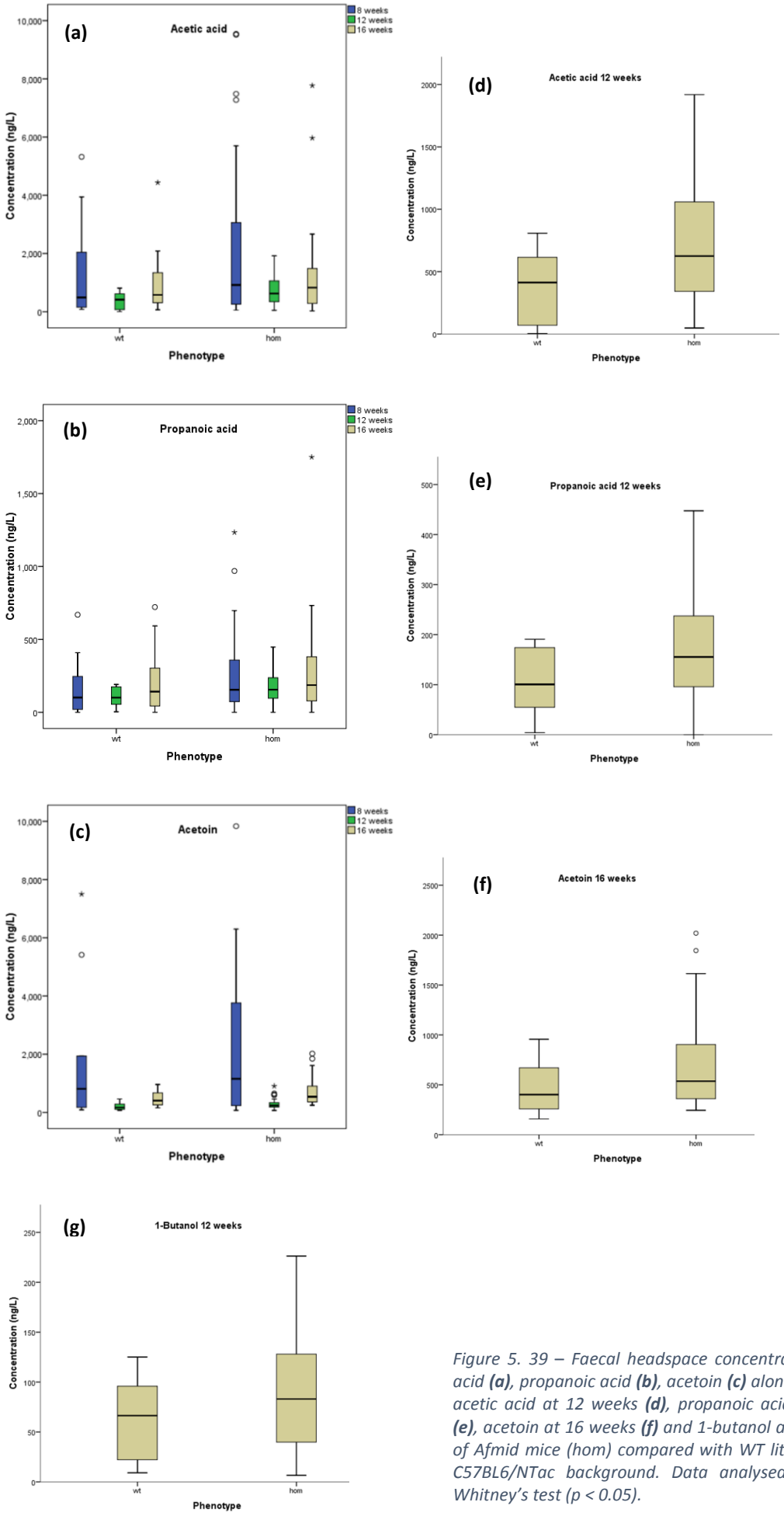


*Figure 5. 39 – Faecal headspace concentration of acetic acid **(a)**, propanoic acid **(b)**, acetoin **(c)** along the age; and acetic acid at 12 weeks **(d)**, propanoic acid at 12 weeks **(e)**, acetoin at 16 weeks **(f)** and 1-butanol at 12 weeks **(g)** of Afmid mice (hom) compared with WT littermates on a C57BL6/NTac background. Data analysed using Mann Whitney's test ($p < 0.05$).*

Friedman test (repeated measures) suggests that are significant differences in acetoin and 1-butanol levels across the age ($p < 0.001$). A Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a statistically significant fluctuation of acetoin with age increase, i.e. from 8 to 12 weeks *($z = -4.68$, $p < 0.001$; median (hom, 8 weeks) = 1.1×10^3/ median (wt, 8 weeks) = 8.1×10^2; median (hom, 12 weeks) = 2.4×10^2/ median (wt, 12 weeks) = 1.7×10^2)*, and from 12 to 16 weeks *($z = -4.94$, $p < 0.001$; median (hom, 16 weeks) = 5.4×10^2/ median (wt, 16 weeks) = 4.0×10^2)*.

A Wilcoxon Signed Rank Test using a Bonferroni adjusted p value established that there is a statistically significant increase of 1-butanol with age increase, specifically between 12 to 16 weeks of age *($z = -3.33$, $p < 0.001$; median (hom, 12 weeks) = 8.3×10^1/ median (wt, 12 weeks) = 6.6×10^1; median (hom, 16 weeks) = 1.5×10^2/ median (wt, 16 weeks) = 8.9×10^1)*.

The relationship between VOC headspace concentration and blood glucose levels was investigated using the non-parametric Spearman's rho correlation coefficient. Outliers have not been removed throughout this study. Two medium correlations indicated by a correlation coefficient between 0.30 to 0.49 ($r = 0.361$, $n = 44$, $p = 0.016$ and $r = 0.357$, $n = 44$, $p = 0.017$) were found between acetoin concentration and blood glucose levels at 12 and 16 weeks of age respectively.

Concluding remarks

Cushing´s mice developed obesity and insulin resistance, whereas *Afmid* mice did not developed obesity and these showed impaired glucose tolerance, although their insulin sensitivity is unchanged once compared to WT littermates.

Univariate analysis and multivariate statistics was performed for the SIFT-MS data, and merely univariate analysis for GC-MS data, since multivariate statistics would require further alignment of the peaks areas prior to statistical analysis. Univariate analysis have identified statistically significant individual *m/z* values, whereas multivariate statistics detected VOC patterns rather promising in separating the groups according to the phenotype.

Overall, faecal headspace analysis by SIFT-MS and GC-MS indicated the presence of several compounds at statistically significant levels such as, short-chain fatty acids (SCFAs), ketones (acetone, 2,3-butanedione, acetoin), alcohols (methanol, ethanol, propanol, 1 butanol) and aldehydes (acetaldehyde).

Significantly increased levels of acetic acid, propanoic acid, butanoic acid and acetone were found for diabetic (het) Cushing´s mice using SIFT-MS, and these findings were validated using GC-MS. Similarly to Cushing´s mice, acetic acid, acetoin and butanoic acid were found to be present in *Afmid* faecal headspace in greater concentration than other compounds by means of using GC-MS. In GC-MS data set, statistically significant elevated VOC concentrations in faecal headspace of acetic acid, propanoic acid, 1-butanol and acetoin were found in mutant (hom) *Afmid* mice, although SIFT-MS results did not confirmed this.

Chapter 6

Analysis of the volatile faecal metabolome in screening for colorectal cancer

6. 1 Introductory notes

There are many potential uses of VOC analysis in medical diagnostics, as discussed in chapter 1. This chapter discusses the potential use of the analysis of the volatile faecal metabolome as a potential screening technique for colorectal cancer.

This work was carried out by a group of colleagues plus myself and published with data analysis carried out by a bioinformatics expert, Dr Michael Cauchi (Batty *et al.*, 2015). The published work may be find in Appendix C.

SIFT-MS measurements were taken and data analysis and statistics displayed in this thesis was carried out by me alone and is shown below. Statistical analysis presented here was performed using the software *IBM SPSS Statistics 21.0,* in contrast to the multivariate statistical techniques used by Cauchi (Batty *et al.*, 2015) which was performed using a higher power computer, which allowed the setup of bootstrapping and testing the stability and reliability of the models, which is able to produce a much more rigorous assessment of statistical significance (Brereton R.G., 2009).

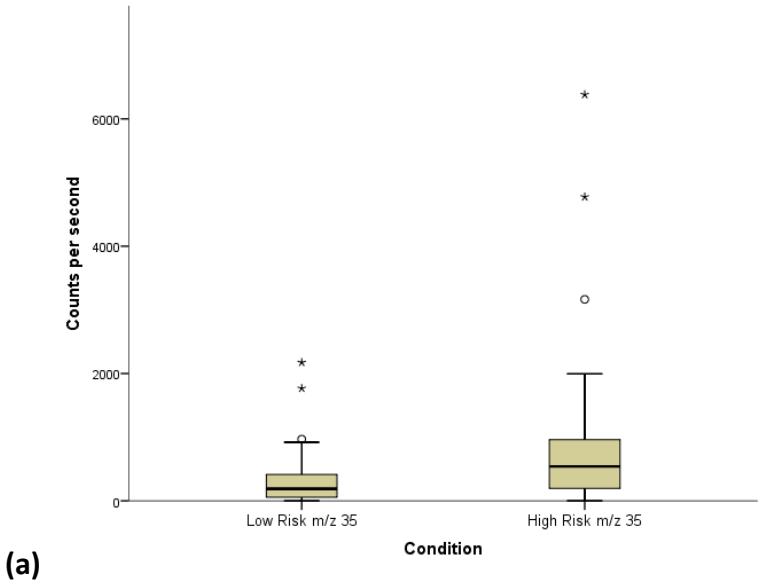
6. 2 Univariate Analysis

The data displayed in this chapter were acquired using the SIFT-MS instrument at The Open University. The normality of the data was assessed. The Kolmogorov-Smirnov statistic and Shapiro-Wilk statistic assessed the normality of the distribution of scores along the data set. Significant results ($p < 0.05$) indicated that the data do not follow normal distribution.

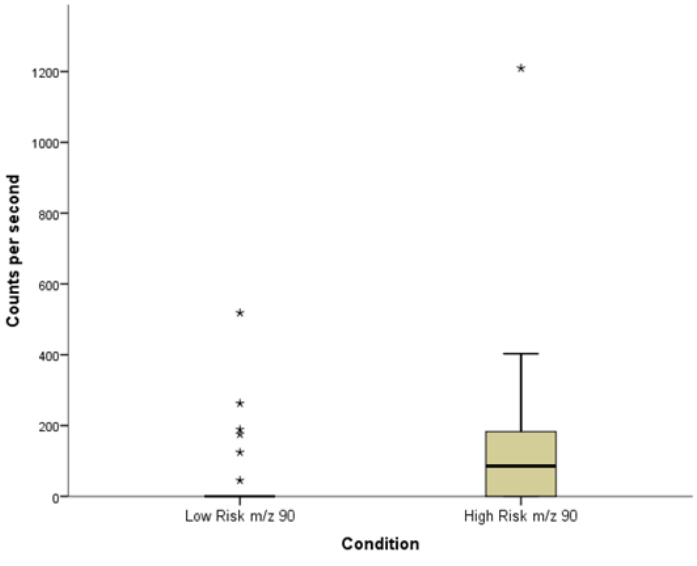
A non-parametric statistic technique, Mann-Whitney U Test, was therefore applied to test the differences between the two independent groups of variables (Low Risk/ High Risk), where Low Risk stands for the 'normal' population and High Risk indicates the group within the class 5 and 6 (high grade adenoma and adenocarcinoma).

The Mann-Whitney U test revealed significant differences in the VOCs levels of low risk and high risk groups, either for H_3O^+, NO^+ and O_2^+ precursor ions. Using H_3O^+ as the precursor ion, statistically significant differences were detected for *m/z* 35, low risk (*median = 188.7, n = 30*) and high risk (*median = 540.8, n = 31*), p = 0.021; and *m/z* 90 where, low risk (*median = 0.0, n =* 30) and high risk (*median = 85.8, n = 31*), p = 0.007.

Figure 6.1 presents the boxplots for *m/z* = 35 (H_3O^+) and *m/z* = 90 (H_3O^+). The boxplots also provide information on outliers represented with a circle and extreme points are indicated with an asterisk.



(a)



(b)

Figure 6. 1 – Boxplots for m/z values showing statistical significance, m/z 35 ***(a)*** *and m/z 90* ***(b)*** *using H_3O^+ precursor ion. Data analysed using Mann-Whitney´s test (p < 0.05).*

A likely identification of *m/z* 35 [H_3S^+] indicates that this VOC might be hydrogen sulphide (Španěl and Smith, 2000b). The potential existence of other overlapping ions at low molecular weight is reduced, thus there is a reasonable probability that in fact this ion is indeed H_2S. Hydrogen sulphide has been identified as a product of microbiota fermentation within the gut (Moore *et al.*, 1987). Sulphate reducing bacteria and existence of hydrogen sulphide in the gut and faeces have been shown to be present in patients with ulcerative colitis and colorectal cancer (Cai *et al.*, 2010, Jia *et al.*, 2012). The unknown variable *m/z* 90 also shown to be statistically significant. The ion *m/z* 90 might possibly be identified as 3-aminopropanoic acid ($C_3H_7NO_2$), commonly known as β-alanine. In literature, this metabolite was found in colon carcinoma tissues with very high statistical significance using GC-TOF (Denkert *et al.*, 2008).

Using NO^+ as precursor ion, the Mann-Whitney U test revealed statistically significant differences for *m/z* 96 [$(CH_3)_2S_2^+$] (Figure 6.2) which may be identified as dimethyl disulphide ($C_2H_6S_2$) (Španěl and Smith, 1998b). Where, low risk (*median = 0.0, n =* 30) and high risk (*median = 58.3, n = 31*), p = 0.026. Although the larger isotope ^{32}S have not shown to be statistically significant at *m/z* 94, the O_2^+ data set revealed that *m/z* 94 was statistically significant between the groups and thus, confirms this *m/z* as dimethyl disulphide.

The presence of sulphides in the gut are controlled by the levels of sulphate reducing bacteria in the gut and these have been shown to be present in patients with colorectal cancer (Gibson *et al.*, 1993, Cai *et al.*, 2010).

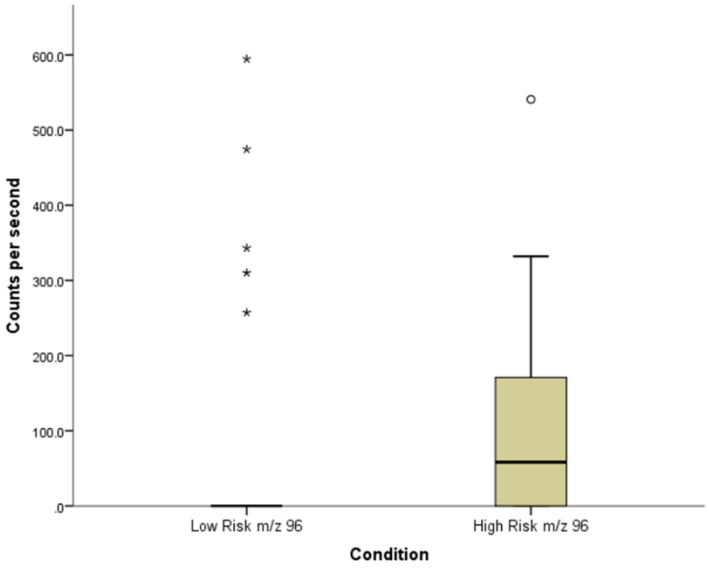


Figure 6. 2 – Boxplots for m/z 96 showing statistical significance, using NO^+ precursor ion. Data analysed using Mann-Whitney´s test (p < 0.05).

Using O_2^+ as precursor ion, statistically significant differences were detected for *m/z* 48, *m/z* 62, *m/z* 66 and *m/z* 94 using Mann-Whitney U test. Median values for each VOC within the Low and High Risk groups are indicated in Table 6.1 with their significance levels.

Table 6. 1 – Median values of Low and High Risk groups for m/z 48, m/z 62, m/z 66 and m/z 94 and its significance levels acquired using O_2^+ precursor ion. Data analysed using Mann Whitney´s test (p < 0.05).

Ion	***Median***		***Significance level (p)***
	Low Risk	**High Risk**	
m/z 48	1533.5	3299.3	0.041
m/z 62	325.6	546.7	0.036
m/z 66	0.0	98.7	0.036
m/z 94	197.2	431.0	0.024

In Figure 6.3 are indicated the boxplots for *m/z* 48, *m/z* 62, *m/z* 66 and *m/z* 94 respectively, acquired using O_2^+ precursor ion.

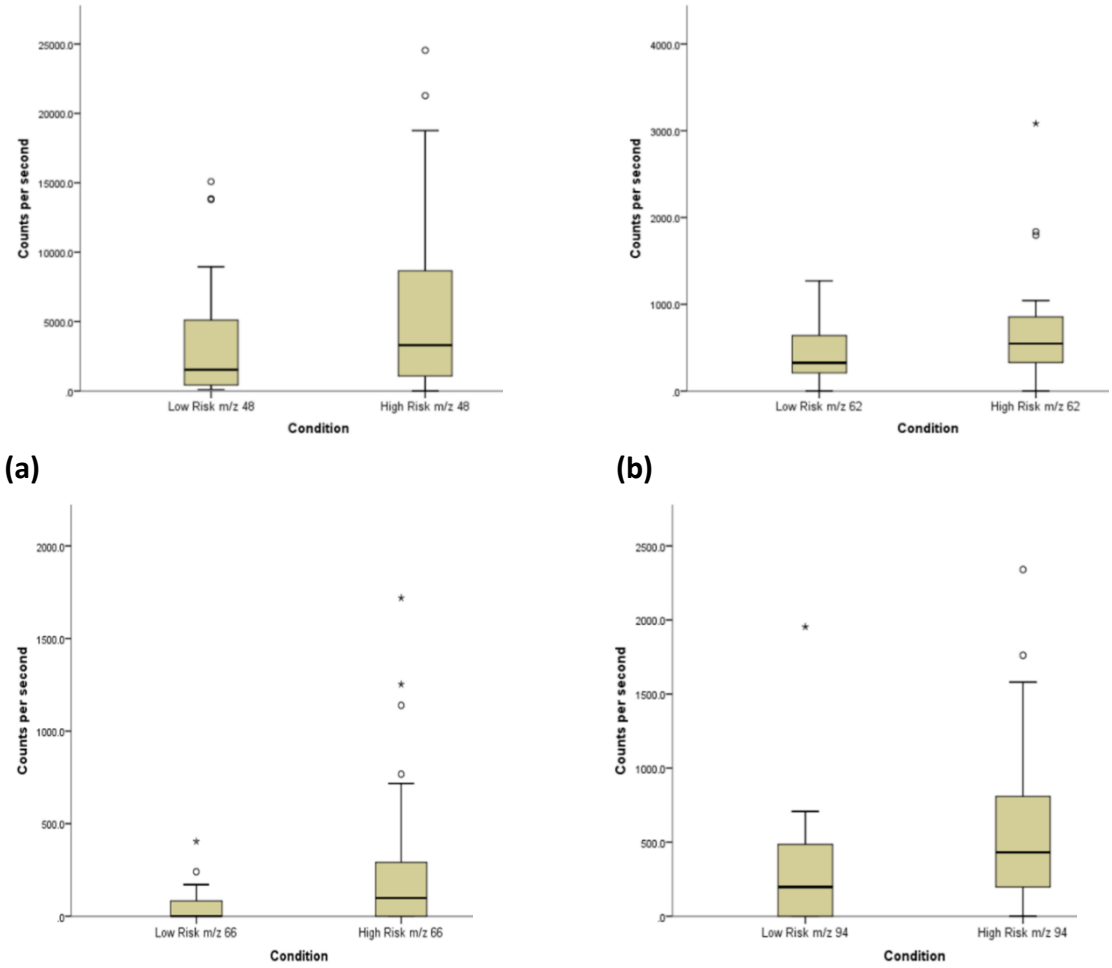


Figure 6. 3 – Boxplots for m/z values showing statistical significance, m/z 48 ***(a)****, m/z 62* ***(b)****, m/z 66* ***(c)****, and m/z 94* ***(d)*** *using O_2^+ precursor ion. Data analysed using Mann-Whitney´s test ($p < 0.05$).*

The ions *m/z* 62 and *m/z* 94 were tentatively identified as dimethyl sulphide [$(CH_3)_2S^+$] and dimethyl disulphide [$(CH_3)_2S_2^+$] (Španěl and Smith, 1998b) respectively, supporting the hypothesis that increased production of sulphides is more predominant in the high risk group. The unknown variables *m/z* 48 and *m/z* 66 are also shown to be statistically significant.

6. 3 Multivariate Statistics

A discriminant analysis was conducted for each data set individually using a single precursor ion, and by combining data sets using H_3O^+, NO^+ and O_2^+ to predict whether a subject was within the Low Risk group or High Risk group.

The cross-validated classification showed an overall good classification for all the precursor ions, and sensitivity and specificity are given (Table 6.2).

Table 6. 2 – Cross-validated classification using Linear Discriminant Analysis and the corresponding test sensitivity and test specificity

Precursor ion	***Cross-validated classification (%)***	***Sensitivity (%)***	***Specificity (%)***
H_3O^+	77.0	68	87
NO^+	68.9	71	67
O_2^+	93.4	90	97
$H_3O^+ + NO^+ + O_2^+$	68.9	67	71

In a national screening programme, although it is highly desirable to detect cancers, it is also very important to minimise the number of false positives obtained in any test. This will reduce the number of unnecessary invasive, unpleasant and expensive diagnostic tests (e.g. colonoscopies). For this reason, it is desirable to have a high specificity in any screening test, in other words to limit the number of false positives. Taking into account this particular requirement, therefore, the best results were achieved using the data sets H_3O^+, O_2^+ and by combining data sets ($H_3O^+ + NO^+ + O_2^+$). The cross-validated classification showed that overall 77.0% were correctly classified using H_3O^+ data set, which had a specificity of 87% and sensitivity of 68%. Although the O_2^+ classification results seem promising, these would need to be further validated with other statistical models.

Using the data set for H_3O^+, the best set of predictors that discriminate between the groups are indicated in Table 6.3. *Wilks´ Lambda* indicates the significance of the discriminant function, thus all predictors are significant ($p < 0.000$).

Table 6. 3 – Best set of predictors discriminating between Low Risk group and High Risk group using the data set H_3O^+

Variables (m/z)	***Wilks' Lambda (p < 0.000)***	***Possible compound (s)***
35	0.000	Hydrogen sulphide
42	0.000	Acetonitrile
77	0.000	Acetone, Carbon disulphide

The metabolic profile given through the multivariate analysis is in agreement with the results assessed using univariate analysis, in which *m/z* 35 was found to be responsible for separating the groups. This leads to the conclusion that *m/z* 35 might be used as clinically relevant biomarker for colorectal cancer. The variable *m/z* 42 is possibly acetonitrile, which has been known to be an exogenous compound. Acetonitrile concentrations were commonly found in the breath and urine of smokers (Abbott *et al.*, 2003). The ion *m/z* 77 might be acetone or carbon disulphide, or a mixture of both.

Significant higher abundances of *Proteobacteria* were found in patients with colorectal adenomas and a significant lower abundance of *Bacteroidetes* (Shen *et al.*, 2010). Thus there is evidence of changes in the gut flora in patients with colorectal cancer. Adenomatous polyps are a precursor form of colorectal cancer, i.e. polyps are not usually cancerous, although it might eventually turn into cancer if left untreated (Leslie *et al.*, 2002). A recent study indicated changes in faecal microbiota associated with adenomatous polyps (Hale *et al.*, 2017).

The results shown in this thesis may indicate a change in the gut microbiota and the existing of sulphate reducing bacteria in the gut which process available hydrogen (Gibson *et al.*, 1993), or in fact that the sulphate reducing bacteria are present in a greater amount in the bowel of the high risk group. Sulphide compounds, including hydrogen sulphide appear to offer potential as biomarkers for screening of colorectal cancer.

Chapter 7

Conclusions and future work

7. 1 Conclusions

This thesis aimed to investigate the use of VOCs in disease diagnosis and monitoring, through the faecal headspace analysis of mouse models of type 2 diabetes, and human studies analysing faecal headspace from patients with colorectal cancer. Faecal headspace studies using mouse models of type 2 diabetes (T2D) have never been performed before. Fundamental chemistry studies have been performed due to the need of having standardised data, and thus investigates a potential method for standardising the analysis of breath gas VOCs.

Chapter 4 discussed a potential method for standardising the analysis of breath gas VOCs using different analytical techniques, in order to get consistent results amongst such techniques.

The results of the present study of the reactions of H_3O^+ with six VOCs biologically significant, comprising several alcohols, one ketone and one aldehyde, indicates that most of the reactions result in multiple products ions, and the abundance and stability of these ions strongly depends on the *E/N* ratio used. It is critical that the reaction rate coefficients for the reactions between H_3O^+ and VOC be determined under the actual conditions in order to obtain reliable quantification.

In contrast to SIFT-MS, in PTR-MS the underlying ion chemistry is often not known, specifically, the kinetics of the ion-molecule reactions and reaction time are not well established and can be very sensitive to changes in *E/N* (Warneke *et al.*, 1996). Thus, careful calibration of the instrument is usually carried out for each VOC and is presently the preferred method to ensure accurate quantification (Beauchamp *et al.*, 2013).

Nevertheless, quantification of VOC concentrations may be accomplished if proton transfer reaction rate coefficients are known (Cappellin *et al.*, 2012).

Product ions and branching ratios were determined at specific working conditions, and subsequently, quantitative determination of VOC concentrations evaluated. Calibration curves determined using SIFT-MS, PTR-MS and TD-GC-MS are given within the range 10^1-10^3 ppbv.

This study demonstrates the potential for overcoming inherent difficulties in the standardisation procedure required for the trace gas analysis of breath constituents. In this work, the findings give slightly different results, but the approach taken of using standard headspaces enables a proper comparison to be made between the three analytical techniques (SIFT-MS, PTR-MS and TD-GC-MS) and enables these techniques to be tested for their accuracy and reproducibility. Additionally this study provides an estimate of how accurately rate coefficients can be determined in SIFT-MS and PTR-MS. This may be one way of standardising methods given that different laboratories use different sampling and analytical techniques. In identifying biomarkers with techniques, such as GC-MS, some compounds cannot be accurately quantified, particularly across a wide range of size and volatility or even polarity. For instance, this study has shown that while SIFT-MS and PTR-MS may be able to detect ethanol and acetaldehyde simply and accurately, a typical TD-GC-MS technique used in VOC biomarker identification is poor at quantification of smaller compounds due to the need to carefully pick sorbents to cover the range.

This work has demonstrated that VOCs may be measured at trace concentrations via different analytical techniques, and the measured levels depend upon several factors, i.e.

collection, storage and analysis. The lack of standardization between analytical techniques is still an issue to deal with, which in turn promotes the existence of inconsistent published results. Testing standards calibrated for the gas phase according to Henry´s law may be an easy option to overcome those difficulties.

Chapter 5 discusses the potential use of VOCs as a potential non-invasive method for monitoring the progress of diabetes. Mice have the potential to provide large data sets for metabolic profiling at controlled conditions. The results shown here were from studying two different mouse models of diabetes, namely, the Cushing´s syndrome mouse model of type 2 diabetes, and single *Afmid* knockout mice exhibiting impaired glucose tolerance.

Type 2 Diabetes (T2D) is a metabolic disease and it has been attributed to changes in diet, poor physical activity, life-style, genetic predisposition, while more recently some studies suggest that there is a link between metabolic diseases and bacterial populations in the gut. Alterations in the gut microbiota have been linked to obesity (Ley *et al.*, 2006), insulin resistance and T2D (Rabot *et al.*, 2010). However, whether gut microbiota plays a causal role in the pathogenesis of metabolic diseases, and the mechanism by which modifications to the gut microbiota might lead to these conditions in humans is unknown. There is evidence of changes in the gut microbiota of type-2 diabetic patients when compared to non-diabetic subjects (Larsen *et al.*, 2010). In this study the composition of faecal microbiota in adults with T2D was compared to non-diabetic controls and the relative abundance of bacterial species *Firmicutes* was significantly lower, while the proportion of *Bacteroidetes* and *Proteobacteria* was somewhat higher in diabetics compared to non-diabetic controls. Accordingly, the ratios of *Bacteroidetes* to *Firmicutes*

are found to be positively correlated with reduced glucose tolerance. Two other studies indicated that patients with type-2 diabetes had less butyrate-producing bacteria than controls (Qin *et al.*, 2012, Karisson *et al.*, 2013). However, it is important to note that studies' discrepancies may be due to ethnic factors, dietary differences, and intake of medication to treat T2D which greatly influences bacterial composition in the gut (Shin *et al.*, 2014).

Gut microbiota are thought to influence glucose and energy metabolism through the production of SCFAs (short-chain fatty acids). These volatile SCFAs include butyric acid also known as butanoic acid, acetic acid, propionic acid also known as propanoic acid, formic acid, isobutyric acid, valeric acid, isovaleric acid, and caproic acids (Bergman, 1990), although acetate, propionate and butyrate resemble 90 to 95% of SCFAs within the gut. These SCFAs are produced through fermentation of undigested dietary carbohydrates by gut bacteria, particularly by *Bacteroides* (Bäckhed *et al.*, 2005), along with the production of CO_2, H_2, and CH_4. The faecal SCFAs concentration in healthy humans has been studied earlier (McOrist *et al.*, 2008) and these concentrations strongly depend on the dietary regime, medication, age, and genetic background.

Alterations in the gut microbiota, and consequently in SCFAs composition have been assumed to be associated with the development of obesity, insulin resistance and diabetes. Mice studies indicated that germ-free mice developed enhanced insulin sensitivity with improved glucose tolerance when fed a high-fat diet (HFD) (Rabot *et al.*, 2010). It was suggested that SCFAs may regulate gut hormones via their endogenous receptors Free fatty acid receptors 2 (FFAR2) and 3 (FFAR3), but direct evidence is lacking. In earlier studies, SCFAs were administrated in mice, and the findings suggested that

butyrate, propionate and acetate protect against diet-induced obesity and insulin resistance (Lin *et al.*, 2012, Gao *et al.*, 2009). Butyrate and propionate, but not acetate, induce gut hormones and reduce food intake (Lin *et al.*, 2012). A hypothetical model for propionic acid effects in humans was also reported (Al-Lahham *et al.*, 2010) suggesting that propionic acid reduces food intake, reduces inflammation and consequently insulin resistance.

Headspace of faecal samples have not been investigated before and the fact that the acids spotted in this thesis are statistically significant may hint at different mechanisms occurring in the gut of *Afmid* and Cushing´s mice.

Importantly, the composition of the gut microbiota are not constant over the age. Thus, the faecal flora in humans of different ages, ranging from 3 to 89 years old, was studied and results indicated that there are significant differences between elderly people and younger adults and children, although the major SCFAs (acetate, propionate and butyrate) within the gut did not suffered alterations across the ageing process (Andrieux *et al.*, 2002).

In the animal work reported in this thesis, Cushing´s mice faecal headspace suffered alterations over the age. Using H_3O^+ data set, significant differences in *m/z* 61 (likely acetic acid according to NO^+ results) across the age ($p < 0.001$) were found, i.e. from 8 to 12 weeks ($z = -4.048$, $p < 0.001$), in Cushing´s mice. In addition, there was a statistically significant increase in *m/z* 65 (possibly ethanol) between 8 to 12 weeks of age, and between 8 to 20 weeks of age. For *Afmid* mice, the faecal headspace suffered alterations over the age, this means that the ion *m/z* 59 (most likely acetone) significantly changed over the age ($p < 0.001$) and specifically between 8 to 12 weeks of age.

Both increases (Li *et al.*, 2014, Rahat-Rozenbloom *et al.*, 2014) and decreases (Murphy *et al.*, 2010) in plasma and faecal SCFAs concentrations were associated with overfeeding, obesity and the metabolic syndrome. A recent and detailed study (Perry *et al.*, 2016) reported changes to gut microbiota, and altered faecal SCFAs concentrations, and these have been associated with obesity, insulin resistance and the metabolic syndrome. Increased production of acetate by an altered gut microbiota in rodents on a high-fat diet (HFD) led to activation of the parasympathetic nervous system – the part of the nervous system that controls 'subconscious' operations such as heart rate and digestion – which in turn, promoted increased glucose output and stimulated the secretion of insulin in rodents, increased ghrelin (hormone produced in the gastrointestinal tract with a significant role in regulating appetite) secretion, abnormally increased appetite, and obesity. They also observed that injections of acetate into rats fed a normal diet stimulated insulin secretion by beta-cells in the pancreas, indicating that acetate was responsible for this effect, however the mechanism is unknown. The relationship between the gut microbiota and increased insulin was studied after transferring faecal matter from chow- or HFD-fed donor rats to chow- or HFD-fed recipients, they reported similar changes in the gut microbiota, acetate levels, and insulin, except for chow-fed donors into chow-fed recipients in which the microbiota did not changed or metabolic phenotypes. Thus, plasma and faecal acetate concentrations were significantly increased in insulin-resistant rats after 3 days or 4 weeks on a HFD diet, which in turn leads to increased food intake, and drives obesity and insulin resistance. These results support the findings presented in this thesis, in which greater concentrations of acetic acid were found for diabetic (het) mice, either by SIFT-MS and GC-MS. The mutant Cushing´s mice developed obesity and showed significantly higher blood glucose concentrations and

plasma insulin concentrations, supporting the hypothesis that Cushing's mice develop a diabetic phenotype related to a gut microbiota interaction.

No significant differences in body weight were detected between *Afmid* mice and WT littermates, thus *Afmid* mice do not develop obesity. *Afmid* mice show impaired glucose tolerance, although their insulin sensitivity is unchanged when compared to WT littermates. Greater concentrations of acetic acid were found for mutant (hom) *Afmid* mice by GC-MS, although SIFT-MS results did not confirm this.

All the animals were fed *ad libitum* on a commercial diet (5.3% fat [corn oil], 21.2% protein, 57.4% carbohydrate, 4.6% fibre; Rat and Mouse Diet No. 3 (RM3). All the animals (het, hom and WT littermates) were placed under the same diet, however we do not know whether diabetic animals ate the same amount as WT littermates.

In this work, the ketones (acetone, 2,3-butanedione, acetoin), alcohols (methanol, ethanol, propanol, 1-butanol) and aldehydes (acetaldehyde) were also found to be statistically significant in the measured samples. Ketones bodies were found to be produced by the liver during fatty-acid and carbohydrate metabolism and related to glucose metabolism in humans (Kalapos, 2003, Decombaz *et al.*, 1983). Interestingly, in this work acetoin (3-hydroxy-2-butanone) was found in higher levels in diabetic mice (het and hom) and there is evidence that acetoin is a product of fermentation in bacteria (Xiao and Xu, 2007, Lopez *et al.*, 1975). *Bacillus subtilis*, *Bacillus amyloliquefaciens*, *Enterobacter cloacae*, *Serratia marcescens*, and *Paenibacillus polymyxa*, can produce acetoin from pyruvate via α-acetolactate by two enzymatic steps catalysed by α-acetolactate synthase

and α-acetolactate decarboxylase (Xiao and Xu, 2007). Acetoin can be further converted to 2,3-butanediol (Bae *et al.*, 2016).

Bacteria present in the gut produce alcohols, including methanol, ethanol, propanol and butanol (Ewen *et al.*, 2005). Long-chain alcohols may be made by reduction of the corresponding acid, therefore butanoic acid may be reduced to 1-butanol (Garner *et al.*, 2007). Acetaldehyde is a toxic intermediate breakdown product of ethanol metabolism by the liver and small amounts of acetaldehyde are produced naturally through gut microbial fermentation and high acetaldehyde levels indicating to promote carcinogenesis in rats (Seitz *et al.*, 1990).

Finger-prick blood tests for monitoring T2D are painful and patients tend to avoid them, either due to the discomfort or phobia. VOC analysis of breath or body fluids may provide non-invasive alternative method for monitoring the disease, however a lot of research is still required. As VOC trials are expensive and difficult in human volunteers, initial studies using mouse models are useful to provide insight about the disease at controlled conditions. The gut microbiota of diabetic mice seem to have a different composition once compared to WT littermates, and fermentation products within the gut seem to appear in different concentrations, including short-chain fatty acids (SCFAs), alcohols, ketones and aldehydes.

Chapter 6 discussed the use of VOCs as a screening tool of colorectal cancer. It has been suggested that colorectal cancer risk is determined by the interaction between diet and microbial metabolism within the gut (O'Keefe *et al.*, 2007). Colorectal cancer risk has been attributed to a diet regime high in red meat and fat leading to protein fermentation

metabolites potentially carcinogenic and likely linked to colon cancer (Hughes *et al.*, 2000, Larsson and Wolk, 2006). Sulphur compounds are generally assimilated in the human diet through protein (Carbonero *et al.*, 2012), thus, sulphate reducing bacteria use these sulphur residues from meat. The results shown in this thesis may indicate a change in the gut microbiota and the existing of sulphate reducing bacteria in the gut which process available hydrogen (Gibson *et al.*, 1993), or in fact that the sulphate reducing bacteria are present in a greater amount in the bowel of the high risk group.

The association between colorectal cancer and gut microbiota has been studied for many years using conventional culture methods and the majority of the VOC-based studies are relatively small. In a recent study the total bacterial counts, in particular anaerobic bacteria, was found to be significantly lower (10.3 ± 0.7 vs. 10.8± 0.3 $\log_{10}$ cells/g of faeces; $p < 0.001$) in patients with colorectal cancer than in healthy individuals (Ohigashi *et al.*, 2013). The concentrations of short-chain fatty acids (SCFAs) produced by microbiota carbohydrate fermentation, such as acetic acid, propionic acid, and butyric acid, were significantly decreased in the diseased group. Significant higher abundances of *Proteobacteria* were found in patients with colorectal adenomas and a significant lower abundance of *Bacteroidetes* (Shen *et al.*, 2010). Thus there is evidence of changes in the gut flora in patients with colorectal cancer. In addition there is increasing evidence that the development of colorectal cancer is promoted when the gastrointestinal/faecal pH is alkaline, which is directly related to the concentrations of organic acids produced by bacteria (Ohigashi *et al.*, 2013, Pye *et al.*, 1990). However, our findings do not provide evidence of increased acid levels.

Following the multivariate analysis on the SIFT-MS data, the analytical method used showed a 77.0% correct classification using H_3O^+ data set. Along with a specificity of 87%

and sensitivity of 68%, meaning that overall this method shows promise as a replacement for the faecal occult blood test (FOBT). Remarkably, through the use of the volatile faecal profile, this study was able to discriminate the groups better than the FOBT in which all the subjects have tested positive for FOBT initially. In order to replace FOBT as a screening test, any replacement would need to be better than FOBT in at least one respect and at least as good in all other respects. It must also be a similar cost, be easy to use and quick to perform, and potentially of greater accuracy. While the present method shows potential it will have to be validated using a much greater number of patients, however, including those with pre-cancerous polyps.

Sulphide compounds, including hydrogen sulphide appear to offer potential as biomarkers for screening of colorectal cancer, although further studies will be required to confirm this.

The published literature in faecal headspace analysis in disease diagnosis is scarce. VOC analysis shows promise for the screening of CRC, non-invasively and painlessly.

7. 2 Future work

Future work comprises the standardization of additional VOCs using different analytical techniques, through the use of standards calibrated for the gas-phase at physiologically representative concentrations.

Additional CRC studies comprising a larger group of patients in order to validate the pilot study presented here as a screening test of CRC; follow-up of the statistically significant compounds (sulphide compounds) identified here, using Multiple Ion Mode (MIM) of SIFT-MS and Selective Ion Monitoring (SIM) mode of GC-MS.

The animal work presented here should be further investigated. Human studies are essential, with testing the breath and urine headspace of many patients at a range of

blood glucose concentrations at many time points. Without this basic research, it is very unlikely that a VOC substitute for direct blood glucose testing will ever be realised.

References

ABBOTT, S. M., ELDER, J. B., ŠPANĚL, P. & SMITH, D. 2003. Quantification of acetonitrile in exhaled breath and urinary headspace using selected ion flow tube mass spectrometry *Int. J. Mass Spectrom.,* 228**,** 655-665.

ADAMS, N. G. & SMITH, D. 1976. The Selected Ion Flow Tube (SIFT): A technique for studying thermal energy ion-neutral reactions. *Int. J. Mass. Spectrom. Ion Physics,* 21**,** 349 - 359.

AGRAWAL, R. P., SHARMA, N., RATHORE, M. S., GUPTA, V. B., JAIN, S., AGARWAL, V. & GOYAL, S. 2013. Noninvasive method for glucose level estimation by saliva. *J. Diabetes Metab.,* 4**,** 1-5.

AIGNER, B., RATHKOLB, B., HERBACH, N., HRABÉ DE ANGELIS, M., WANKE, R. & WOLF, E. 2008. Diabetes models by screen for hyperglycemia in phenotype-driven ENU mouse mutagenesis projects. *Am. J. Physiol. Endocrinol. Metab.,* 294**,** E232-E240.

AJIBOLA, O. A., SMITH, D., ŠPANĚL, P. & FERNS, G. 2013. Effects of dietary nutrients on volatile breath metabolites. *J. Nutr. Sci.,* 2**,** 1-15.

AL-LAHHAM, S. H., PEPPELENBOSCH, M. P., ROELOFSEN, H., VONK, R. J. & VENEMA, K. 2010. Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. *Biochim. Biophys. Acta,* 1801**,** 1175-1183.

ALASALVAR, C., GRIGOR, J. M., ZHANG, D., QUANTICK, P. C. & SHAHIDI, F. 2001. Comparison of volatiles, phenolics, sugars, antioxidant vitamins, and sensory quality of different colored carrot varieties. *Journal of Agricultural and Food Chemistry,* 49**,** 1410-1416.

ALLISON, J. E., TEKAWA, I. S., RANSOM, L. J. & ADRAIN, A. L. 1996. A comparison of fecal occult-blood tests for colorectal-cancer screening. *N. Engl. J. Med.,* 334(3)**,** 155-159.

ALPHATRAK2. Product information. Available: https://www.zoetisus.com/contact/pages/product_information/msds_pi/pi/alphatrak_2.pdf.

ALTOMARE, D. F., LENA, M. D., PORCELLI, F., TRIZIO, L., TRAVAGLIO, E., TUTINO, M., DRAGONIERI, S., MEMEO, V. & GENNARO, G. 2013. Exhaled volatile organic compounds identify patients with colorectal cancer. *Brit. J. Surg.,* 100**,** 144-150.

AMANN, A., MIEKISCH, W., PLEIL, J., RISBY, T. & SCHUBERT, J. 2010. Methodological issues of sample collection and analysis of exhaled breath. *Exhaled Biomarkers*, 49, 96-114.

AMANN, A. & SMITH, D. 2013. *Volatile Biomarkers Non-Invasive Diagnosis in Physiology and Medicine,* Oxford, Elsevier.

ANDERSON, J., BABB, A. & HLASTALA, M. P. 2003. Modeling soluble gas exchange in the airways and alveoli. *Ann. Biomed. Eng.,* 31**,** 1402-1422.

ANDERSON, J. & HLASTALA, M. 2007. Breath tests and airway gas exchange. *Pulm. Pharmacol. Ther.,* 20**,** 112-117.

ANDERSON, J., LAMM, W. & HLASTALA, M. 2006. Measuring airway exchange of endogenous acetone using a single-exhalation breathing maneuver. *J. Appl. Physiol.,* 100**,** 880-889.

ANDRIEUX, C., MEMBRÉ, J. M., CAYUELA, C. & ANTOINE, J. M. 2002. Metabolic characteristics of the faecal microflora in humans from three age groups. *Scand J Gastroenterol.,* 37**,** 792-798.

ASHCROFT, A. E. & BARNETT, N. W. 1997. *Ionization methods in organic mass spectrometry,* Cambridge, The Royal Society of Chemistry.

BÄCKHED, F., RUTH E. LEY, SONNENBURG, J. L., PETERSON, D. A. & GORDON, J. I. 2005. Host-bacterial mutualism in the human intestine. *Science,* 307**,** 1915-1920.

BAE, S.-J., KIM, S. & HAHN, J.-S. 2016. Efficient production of acetoin in *Saccharomyces cerevisiae* by disruption of 2,3-butanediol dehydrogenase and expression of NADH oxidase. *Sci. Rep.,* 6**,** 27667.

BATTY, C. A., CAUCHI, M., LOURENÇO, C., HUNTER, J. O. & TURNER, C. 2015. Use of the analysis of the volatile faecal metabolome in screening for colorectal cancer. *PLOS ONE,* 10(6)**,** e0130301.

BEAUCHAMP, J. 2011. Inhaled today, not gone tomorrow: pharmacokinetics and environmental exposure of volatiles in exhaled breath. *J. Breath Res.,* 5**,** 1-14.

BEAUCHAMP, J., HERBIG, J., DUNKL, J., SINGER, W. & HANSEL, A. 2013. On the performance of proton-transfer-reaction mass spectrometry for breath-relevant gas matrices. *Meas. Sci. Technol.,* 24**,** 125003.

BEAUCHAMP, J., HERBIG, J. & GUTMANN, R. 2008. On use of Tedlar bags for breath-gas sampling and analysis. *J. Breath Res.,* 2**,** 1-18.

BEAUCHAMP, J. & PLEIL, J. 2013. Simply breath-taking? Developing a strategy for consistent breath sampling. *J. Breath Res.,* 7**,** 1-3.

BENTLEY, L., ESAPA, C. T., NESBIT, M. A., HEAD, R. A., EVANS, H., LATH, D., SCUDAMORE, C. L., HOUGH, T. A., PODRINI, C., HANNAN, F. M., FRASER, W. D., CROUCHER, P. I., BROWN, M. A., BROWN, S. D. M., COX, R. D. & THAKKER, R. V. 2014. An N-ethyl-N-nitrosourea induced corticotropin releasing hormone promoter mutation provides a mouse model for endogenous glucocorticoid excess. *Endocrinology,* 155**,** 908-922.

BERGMAN, E. N. 1990. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiol. Rev,* 70**,** 567-590.

BIRKEN, T., SCHUBERT, J., MIEKISCH, W. & NÖLDGE-SCHOMBURG, G. 2006. A novel visually CO_2 controlled alveolar breath sampling technique *Technol. Health Care,* 14**,** 499-506.

BLAKE, R. S., MONKS, P. S. & ELLIS, A. M. 2009. Proton-transfer reaction mass spectrometry. *Chem. Rev.,* 109**,** 861-896.

BOHME, D. K., MACKAY, G. I. & SCHIFF, H. I. 1980. Determination of proton affinities from the kinetics of proton transfer reactions. VII. The proton affinities of O_2, H_2, Kr, O, N_2, Xe, CO_2, CH_4, N_2O, and CO. *J. Chem. Phys.,* 73 (10)**,** 4976-4986.

BORSDORF, H. & EICEMAN, G. A. 2006. Ion Mobility Spectrometry: Principles and Applications. *Appl. Spec. Rev.,* 41**,** 323-375.

BRERETON R.G. 2009. *Chemometrics for pattern recognition,* Chichester, John Wiley & Sons.

BRETTHAUER, M. 2011. Colorectal cancer screening. *Intern. Med.,* 270**,** 87-98.

BROWN, P., WATTS, P., MÄRK, T. D. & MAYHEW, C. A. 2010. Proton transfer reaction mass spectrometry investigations on the effects of reduced electric field and reagent ion internal energy on product ion branching ratios for a series of saturated alcohols. *Int. J. Mass Spectrom.,* 294**,** 103-111.

BUSZEWSKI, B., KESY, M., LIGOR, T. & AMANN, A. 2007. Human exhaled air analytics: biomarkers of diseases. *Biomed. Chromatogr.,* 21**,** 553 - 566.

BUSZEWSKI, B., ULANOWSKA, A., LIGOR, T., DENDERZ, N. & AMANN, A. 2009. Analysis of exhaled breath from smokers, passive smokers and non-smokers by solid-phase microextraction gas chromatography/mass spectrometry. *Biomed. Chromatogr.,* 23**,** 551-556.

CAI, W. J., WANG, M. J., LU, L. H., WANG, C. & ZHU, Y. C. 2010. Hydrogen sulfide induces human colon cancer cell proliferation: role of Akt, ERK and p21. *Cell Biol. Int.,* 34**,** 565-572.

CAPPELLIN, L., KARL, T., PROBST, M., ISMAILOVA, O., WINKLER, P. M., SOUKOULIS, C., APREA, E., MÄRK, T. D., GASPERI, F. & BIASIOLI, F. 2012. On quantitative determination of volatile organic compound concentrations using proton transfer reaction time-of-flight mass spectrometry. *Environ. Sci. Technol.,* 46**,** 2283-2290.

CAPPELLIN, L., PROBST, M., LIMTRAKUL, J., BIASIOLI, F., SCHUHFRIED, E., SOUKOULIS, C., MÄRK, T. D. & GASPERI, F. 2010. Proton transfer reaction rate coefficients between H_3O^+ and some sulphur compounds. *Int. J. Mass Spectrom.,* 295**,** 43-48.

CHEVERUD, J. M., EHRICH, T. H., KENNEY, J. P., PLETSCHER, L. S. & SEMENKOVICH, C. F. 2004. Genetic evidence for discordance between obesity- and diabetes- related traits in the LGXSM recombinant inbred mouse strains. *Diabetes,* 53**,** 2700-2708.

CROFFORD, O. B., MALLARD, R. E., WINTON, R. E., ROGERS, N. L., JACKSON, J. C. & KELLER, U. 1977. Acetone in breath and blood. *Trans. Am. Clin. Climatol. Assoc.,* 88**,** 128-139.

CUMMINGS, J. H. & BINGHAM, S. A. 1987. Dietary fibre, fermentation and large bowel cancer. *Cancer Surv.,* 6**,** 601-621.

DASH, S., MUELLER, T. J., VENKATARAMANAN, K. P., PAPOUTSAKIS, E. T. & MARANAS, C. D. 2014. Capturing the response of *Clostridium acetobutylicumto* chemical stressors using a

regulated genome-scale metabolic model. *Biotechnol. Biofuels,* 7**,** doi:10.1186/s13068-014-0144-4.

DE LACY COSTELLO, B., AMANN, A., AL-KATEB, H., FLYNN, C., FILIPIAK, W., KHALID, T., OSBORNE, D. & RATCLIFFE, N. M. 2014. A review of the volatiles from the healthy human body. *J. Breath Res.,* 8**,** 014001.

DE LACY COSTELLO, B., EWER, A. K., GARNER, C. E., PROBERT, C. S., RATCLIFFE, N. M. & SMITH, S. 2008. An analysis of volatiles in the headspace of the faeces of neonates *J. Breath Res.,* 2**,** 1-8.

DECOMBAZ, J., ARNAUD, M. J., MILON, H., MOESCH, H., PHILIPPOSSIAN, G., THELIN, A. L. & HOWALD, H. 1983. Energy metabolism of medium-chain triglycerides versus carbohydrates during exercise. *Eur. J. Appl. Physiol. Occup. Physiol. ,* 52**,** 9-14.

DENG, C., ZHANG, J., YU, X., ZHANG, W. & ZHANG, X. 2004. Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization. *J. Chromatogr. B,* 810**,** 269-275.

DENKERT, C., BUDCZIES, J., WEICHERT, W., WOHLGEMUTH, G., SCHOLZ, M., KIND, T., NIESPOREK, S., NOSKE, A., BUCKENDAHL, A., DIETEL, M. & FIEHN, O. 2008. Metabolite profiling of human colon carcinoma – deregulation of TCA cycle and amino acid turnover. *Mol. Cancer,* 7**,** DOI: 10.1186/1476-4598-7-72.

DI FRANCESCO, F., LOCCIONI, C., FIORAVANTI, M., RUSSO, A., PIOGGIA, G., FERRO, M., ROEHRER, I., TABUCCHI, S. & ONOR, M. 2008. Implementation of Fowler's method for end-tidal air sampling. *J. Breath Res.,* 2**,** 1-9.

DICKERSON, S. S. & KEMENY, M. 2004. Acute stressor and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol. Bull.,* 130**,** 355-391.

DISKIN, A. M., ŠPANĚL, P. & SMITH, D. 2003. Increase of acetone and ammonia in urine headspace and breath during ovulation quantified using selected ion flow tube mass spectrometry. *Physiol. Meas.,* 24**,** 191-199.

DOBROVOLSKY, V. N., BOWYER, J. F., PABARCUS, M. K., HEFLICH, R. H., WILLIAMS, L. D., DOERGE, D. R., ARVIDSSON, B., BERGQUIST, J. & CASIDA, J. E. 2005. Effect of arylformamidase (kynurenine formamidase) gene inactivation in mice on enzymatic activity, kynurenine pathway metabolites and phenotype. *BBA-Gen subjects,* 1724**,** 163-172.

DUBOWSKI, K. M. 1974. Biological aspects of breath-alcohol analysis. *Clin. Chem.,* 20**,** 294-299.

ELLIS, A. M. & MAYHEW, C. A. 2014. *Proton transfer reaction mass spectrometry-Principles and applications,* Chichester, John Wiley & Sons Ltd.

ENDERBY, B., LENNEY, W., BRADY, M., EMMETT, C., ŠPANĚL, P. & SMITH, D. 2009. Concentrations of some metabolites in the breath of healthy children aged 7–18 years measured using selected ion flow tube mass spectrometry (SIFT-MS). *J. Breath Res.,* 3**,** 036001.

EWEN, R., DE LACY COSTELLO, B., GARNER, C., RATCLIFFE, N. M., SMITH, S. & PROBERT, S. J. 2005. Rapid diagnosis of gastro-intestinal infection using faecal odour. *Breath Analysis for Clinical Diagnosis and Therapeutic Monitoring.* Singapore: World Scientific.

FALL, R., KARL, T., JORDAN, A. & LINDINGER, W. 2001. Biogenic C5 VOCs: release from leaves after freeze–thaw wounding and occurrence in air at a high mountain observatory. *Atmos. Environ.,* 35**,** 3905-3916.

FIELD, A. 2013. *Discovering statistics using IBM SPSS statistics,* London, SAGE publications.

FILIPIAK, W., FILIPIAK, A., AGER, C., WIESENHOFER, H. & AMANN, A. 2012. Optimization of sampling parameters for collection and preconcentration of alveolar air by needle traps. *J. Breath Res.,* 6**,** 1-19.

FILIPIAK, W., SPONRING, A., FILIPIAK, A., AGER, C., SCHUBERT, J., MIEKISCH, W., AMANN, A. & TROPPMAIR, J. 2010. TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells *in vitro*. *Cancer Epidemiol Biomarkers Prev,* 19**,** 182-195.

FISCHER, S., TREFZ, P., BERGMANN, A., STEFFENS, M., ZILLER, M., MIEKISCH, W., SCHUBERT, J. S., KÖHLER, H. & REINHOLD, P. 2015. Physiological variability in volatile organic compounds (VOCs) in exhaled breath and released from faeces due to nutrition and somatic growth in a standardized caprine animal model. *J. Breath Res.,* 9**,** 1-17.

FOWLER, W. S. 1948. Lung function studies: II. The respiratory dead space. *Am. J. Physiol.,* 154**,** 405-416.

GAO, Z., YIN, J., ZHANG, J., WARD, R. E., MARTIN, R. J., LEFEVRE, M., CEFALU, W. T. & YE, J. 2009. Butyrate improves insulin sensitivity and increases energy expenditure in mice. *Diabetes,* 58**,** 1509-1517.

GARNER, C. E., SMITH, S., DE LACY COSTELLO, B., WHITE, P., SPENCER, R., PROBERT, C. S. J. & RATCLIFFE, N. M. 2007. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J.,* 21**,** 1675-1688.

GHIMENTI, S., LOMONACO, T., BELLAGAMBI, F. G., TABUCCHI, S., ONOR, M., TRIVELLA, M. G., CECCARINI, A., FUOCO, R. & FRANCESCO, F. D. 2015a. Comparison of sampling bags for the analysis of volatile organic compounds in breath. *J. Breath Res.,* 9**,** 1-13.

GHIMENTI, S., TABUCCHI, S., BELLAGAMBI, F. G., LOMONACO, T., ONOR, M., TRIVELLA, M. G., FUOCO, R. & FRANCESCO, F. D. 2015b. Determination of sevoflurane and isopropyl alcohol in exhaled breathby thermal desorption gas chromatography–mass spectrometry for exposure assessment of hospital staff. *J. Pharm. Biomed. Anal.,* 106**,** 218-223.

GIBSON, G. R., MACFARLANE, G. T. & CUMMINGS, J. H. 1993. Sulphate reducing bacteria and hydrogen metabolism in the human large intestine. *Gut,* 34**,** 437-439.

GILCHRIST, F. J., RAZAVI, C., WEBB, A. K., JONES, A. M., ŠPANĚL, P., SMITH, D. & LENNEY, W. 2012. An investigation of suitable bag materials for the collection and storage of breath samples containing hydrogen cyanide. *J. Breath Res.,* 6**,** 1-7.

GREGER, R. & WINDHORST, U. 1996. *Comprehensive Human Physiology,* Berlin, Springer-Verlag.

GROUP, U. C. C. S. P. 2004. Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom *BMJ.*

GUO, D., ZHANG, D., LI, N., ZHANG, L. & YANG, J. 2010. *Diabetes identification and classification by means of a breath analysis system,* Berlin, Springer.

HABIB, M. P., CLEMENTS, N. C. & GAREWAL, H. S. 1995. Cigarette smoking and ethane exhalation in humans. *Am. J. Respir. Crit. Care Med.,* 151**,** 1368-1372.

HALBRITTER, S., FEDRIGO, M., HOLLRIEGL, V., SZYMCZAK, W., MAIER, J. M., ZIEGLER, A. G. & HUMMEL, M. 2012. Human Breath Gas Analysis in the Screening of Gestational Diabetes Mellitus. *Diabetes Technol. Therapeutics,* 14**,** 917-925.

HALE, V. L., CHEN, J., JOHNSON, S., HARRINGTON, S. C., YAB, T. C., SMYRK, T. C., NELSON, H., BOARDMAN, L. A., DRULINER, B. R., LEVIN, T. R., REX, D. K., AHNEN, D. J., LANCE, P., AHLQUIST, D. A. & CHIA, N. 2017. Shifts in the fecal microbiota associated with adenomatous polyps. *Cancer Epidemiol Biomarkers Prev,* 26**,** 85-94.

HAMILTON, J. G. 1995. Needle phobia - a neglected diagnosis. *J. Fam. Pract.,* 41**,** 169-175.

HELLER, A. 1996. Amperometric sensors. *Curr. Opin. Chem. Biol.,* 7**,** 50-54.

HERBIG, J. & AMANN, A. 2009. Proton Transfer Reaction-Mass Spectrometry Applications in Medical Research. *J. Breath Res.,* 3**,** 1-2.

HERBIG, J. & BEAUCHAMP, J. 2014. Towards standardization in the analysis of breath gas volatiles. *J. Breath Res.,* 8**,** 037101.

HERBIG, J., MÜLLER, M., SCHALLHART, S., TITZMANN, T., GRAUS, M. & HANSEL, A. 2009. On-line breath analysis with PTR-TOF. *J. Breath Res.,* 3**,** 027004.

HERBIG, J., TITZMANN, T., BEAUCHAMP, J., KOHL, I. & HANSEL, A. 2008. Buffered end-tidal (BET) sampling - a novel method for real-time breath-gas analysis. *J. Breath Res.,* 2**,** 1-9.

HEWITT, C. N., HAYWARD, S. & TANI, A. 2003. The application of proton transfer reaction-mass spectrometry (PTR-MS) to the monitoring and analysis of volatile organic compounds in the atmosphere. *J. Environ. Monit.,* 5**,** 1-7.

HITOTSUMACHI, S., CARPENTER, D. A. & RUSSELL, W. L. 1985. Dose-repetition increases the mutagenic effectiveness of *N*-ethyl-*N*-nitrosourea in mouse spermatogonia. *Proc. Natl. Acad. Sci. USA,* 82**,** 6619-6621.

HOFFMAN, E. D. & STROOBANT, V. 2007. *Mass spectrometry, principles and applications.,* Chichester, John Wiley & Sons Ltd.

HOLT, R. I. G. & HANLEY, N. A. 2012. *Essential endocrinology and diabetes,* Oxford.

HORVÁTH, I., HUNT, J. & BARNES, P. J. 2005. Exhaled breath condensate: methodological recommendations and unresolved questions. *Eur. Respir. J.,* 26**,** 523-548.

HUANG, J., KUMAR, S., ABBASSI-GHADI, N., ŠPANĚL, P., SMITH, D. & HANNA, G. B. 2013. Selected ion flow tube mass spectrometry analysis of volatile metabolites in urine headspace for the profiling of gastro-esophageal cancer. *Anal. Chem.,* 85**,** 3409-3416.

HÜBSCHMANN, H. J. 2009. *Handbook of GC/MS: Fundamentals and Applications,* Weinheim, WILEY-VCH Verlag GmbH & Co.

HUGHES, R., MAGEE, E. A. M. & BINGHAM, S. A. 2000. Protein degradation in the large intestine: relevance to colorectal cancer. *Curr. Issues Intest. Microbiol.,* 1**,** 51-58.

HUGILL, A. J., STEWART, M. E., YON, M. A., PROBERT, F., COX, I. J., HOUGH, T. A., SCUDAMORE, C. L., BENTLEY, L., WALL, G., WELLS, S. E. & COX, R. D. 2015. Loss of arylformamidase with reduced thymidine kinase expression leads to impaired glucose tolerance. *Biol. Open,* 4**,** 1367-1375.

HUI, Y. H. 2010. *Handbook of fruit and vegetable flavors,* Hoboken, NJ, USA, John Wiley & Sons.

HUNTER, E. P. & LIAS, S. G. 1998. Evaluated gas phase basicities and proton affinities of molecules: an update. *J. Phys. Chem. Ref. Data,* 27**,** 413-656.

JIA, W., WHITEHEAD, R., GRIFFITHS, L., BAI, H., WARING, R. & RAMSDEN, D. 2012. Diversity and distribution of sulphate-reducing bacteria in human faeces from healthy subjects and patients with inflammatory bowel disease. *FEMS Immunol. Med. Microbiol.,* 65**,** 55-68.

JORDAN, A., HAIDACHER, S., HANEL, G., HARTUNGEN, E., HERBIG, J., MÄRK, L., SCHOTTKOWSKY, R., SEEHAUSER, H., SULZER, P. & MÄRK, T. D. 2009a. An online ultra-high sensitivity proton-transfer-reaction mass-spectrometer combined with switchable reagent ion capability (PTR+SRI-MS). *Int. J. Mass Spectrom.,* 286**,** 32-38.

JORDAN, A., HAIDACHER, S., HANEL, G., HARTUNGEN, E., MÄRK, L., SEEHAUSER, H., SCHOTTKOWSKY, R., SULZER, P. & MÄRK, T. D. 2009b. A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS). *Int. J. Mass Spectrom.,* 286**,** 122-128.

JORDAN, A., HANSEL, A., HOLZINGER, R. & LINDINGER, W. 1995. Acetonitrile and benzene in the breath of smokers and non-smokers investigated by proton transfer reaction mass spectrometry (PTR-MS). *Int. J. Mass Spectrom.,* 148**,** L1-L3.

KALAPOS, M. P. 2003. On the mammalian acetone metabolism: from chemistry to clinical implications. *Biochim. Biophys. Acta,* 1621**,** 122-139.

KARISSON, F. H., TREMAROLI, V., NOOKAEW, I., BERGSTROM, G., BEHRE, C. J., FAGERBERG, B., NIELSEN, J. & BACKHED, F. 2013. Gut metagenome in European women with normal, impared and diabetic glucose control. *Nature,* 498**,** 99-103.

KARL, T., PRAZELLER, P., MAYR, D., JORDAN, A., RIEDER, J., FALL, R. & LINDINGER, W. 2001. Human breath isoprene and its relation to blood cholesterol levels: new measurements and modeling. *J. Appl. Physiol.,* 91**,** 762-770.

KEBARLE, P. & GODBOLE, E. W. 1963. Mass-spectrometric study of ions from alpha-particle irradiation of gases at near atmospheric pressures. *J. Chem. Phys.,* 39**,** 1131.

KHALID, T. Y., COSTELLO, B. L., EWEN, R., WHITE, P., STEVENS, S., GORDON, F., COLLINS, P., MCCUNE, A., SHENOY, A., SHETTY, S., RATCLIFFE, N. M. & PROBERT, C. S. 2013a. Breath volatile analysis from patients diagnosed with harmful drinking, cirrhosis and hepatic encephalopathy: a pilot study. *Metabolomics,* 9**,** 938-948.

KHALID, T. Y., SAAD, S., GREENMAN, J., COSTELLO, B. L., PROBERT, C. S. J. & RATCLIFFE, N. M. 2013b. Volatiles from oral anaerobes confounding breath biomarker discovery. *J. Breath Res.,* 7**,** 1-12.

KING, J., KOC, H., UNTERKOFLER, K., MOCHALSKI, P., KUPFERTHALER, A., TESCHL, G., TESCHL, S., HINTERHUBER, H. & AMANN, A. 2010. Physiological modeling of isoprene dynamics in exhaled breath. *J. Theor. Biol.,* 267**,** 626-637.

KING, J., UNTERKOFLER, K., TESCHL, G., TESCHL, S., KOC, H., HINTERHUBER, H. & AMANN, A. 2011. A mathematical model for breath gas analysis of volatile organic compounds with special emphasis on acetone. *J. Math. Biol.,* 63**,** 959-999.

KING, J., UNTERKOFLER, K., TESCHL, S., MOCHALSKI, P., KOC¸, H., HINTERHUBER, H. & AMANN, A. 2012. A modeling-based evaluation of isothermal rebreathing for breath gas analyses of highly soluble volatile organic compounds. *J. Breath Res.,* 6**,** 1-10.

KING, J., UNTERKOFLER, K., TESCHL, S., S TESCHL4, MOCHALSKI, P., KOC¸, H., HINTERHUBER, H. & AMANN, A. 2009. Isoprene and acetone concentration profiles during exercise on an ergometer. *J. Breath Res.,* 6**,** 1-11.

KISTLER, M., MUNTEAN, A., SZYMCZAK, W., RINK, N., FUCHS, M., GAILUS-DURNER, V., WURST, W., HOESCHEN, C., KLINGENSPOR, M., ANGELIS, M. H. & ROZMAN, J. 2016. Diet-induced and mono-genetic obesity alter volatile organic compound signature in mice. *J. Breath Res.,* 10**,** 1-15.

KISTLER, M., SZYMCZAK, W., FEDRIGO, M., FIAMONCINI, J., HÖLLRIEGL, V., HOESCHEN, C., KLINGENSPOR, M., ANGELIS, M. H. & ROZMAN, J. 2014. Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice. *J. Breath Res.,* 8**,** 1-9.

KORE TECHNOLOGY, L. Manual Z-5851-M.

KUFE, D. W., POLLOCK, R. E., WEICHSELBAUM, R. R., BAST JR., R. C., GANSLER, T. S., HOLLAND, J. F. & FREI, E. 2003. *Cancer medicine,* People´s medical publishing house-USA, B C Decker Inc.

KUSHCH, I., SCHWARZ, K., SCHWENTNER, L., BAUMANN, B., DZIEN, A., SCHMID, A., UNTERKOFLER, K., GASTL, G., ŠPANĚL, P., SMITH, D. & AMANN, A. 2008. Compounds enhanced in a mass spectrometric profile of smokers' exhaled breath versus non-smokers as determined in a pilot study using PTR-MS. *J. Breath Res.,* 2**,** 1-26.

LALKHEN, A. G. & MCCLUSKEY, A. 2008. Clinical tests: sensitivity and specificity. *Contin. Educ. Anaesth. Crit. Care Pain,* 8**,** 221-223.

LANGEVIN, P. 1905. A fundamental formula of kinetic theory. *Ann. Chim. Phys.,* 5**,** 245-288.

LARSEN, N., VOGENSEN, F. K., VAN DEN BERG, F. W. J., NIELSEN, D. S., S., A. A., PEDERSEN, B. K., AL-SOUD, W. A., SØRENSEN, S. J., HANSEN, L. H. & JAKOBSEN, M. 2010. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLOS ONE,* 5**,** e9085.

LARSSON, S. & WOLK, A. 2006. Meat consumption and risk of colorectal cancer: A meta-analysis of prospective studies. *Int. J. Cancer,* 119**,** 2657-2664.

LECHNER, M., MOSER, B., NIEDERSEER, D., KARLSEDER, A., HOLZKNECHT, B., FUCHS, M., COLVIN, S., TILG, H. & RIEDER, J. 2006. Gender and age specific differences in exhaled isoprene levels. *J. Respir. Physiol. Neurobiol. ,* 154**,** 478-483.

LEFEVRE, F. & ARONSON, N. 2000. Ketogenic diet for the treatment of refractory epilepsy in children: a systematic review of efficacy. *Pediatrics,* 105**,** 1-7.

LESLIE, A., CAREY, F. A., PRATT, N. R. & STEELE, R. J. C. 2002. The colorectal adenoma–carcinoma sequence. *Br J Surg,* 89**,** 845–860.

LEVITT, M., ELLIS, C. & FURNE, J. 1998. Influence of method of alveolar air collection on results of breath tests. *Dig. Dis. Sci.,* 43**,** 1938-1945.

LEY, R. E., TURNBAUGH, P. J., KLEIN, S. & GORDON, J. I. 2006. Microbial ecology: human gut microbes associated with obesity. *Nature,* 444**,** 1022-1023.

LI, M., GU, D., XU, N., LEI, F., DU, L., ZHANG, Y. & XIE, W. 2014. Gut carbohydrate metabolism instead of fat metabolism regulated by gut microbes mediates high-fat diet-induced obesity. *Benef. Microbes,* 5**,** 335-344.

LIAS, S. G., BARTMESS, J. E., LIEBMAN, J. F., HOLMES, J. L., LEVIN, R. D. & MALLARD, W. G. 1988. Gas-phase ion and neutral thermochemistry. *j. Phys. Chem. Ref. Data,* 17**,** 1-861.

LIDE, D. R. 2014. *Handbook of Chemistry and Physics* CRC Press.

LIKHODII, S., MUSA, K. & CUNNANE, S. 2002. Breath acetone as a measure of systemic ketosis assessed in a rat model of the ketogenic diet. *Clin. Chem.,* 48**,** 115-120.

LIMA, P. O., CALIL, C. M. & MARCONDES, F. K. 2013. Influence of gender and stress on the volatile sulfur compounds and stress biomarkers production. *Oral dis.,* 19**,** 366-373.

LIN, H. V., FRASSETTO, A., KOWALIK JR, E. J., NAWROCKI, A. R., LU, M. M., KOSINSKI, J. R., HUBERT, J. A., SZETO, D., YAO, X., FORREST, G. & MARSH, D. J. 2012. Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PLOS ONE,* 7**,** e35240.

LINDINGER, C., POLLIEN, P., ALI, S., YERETZIAN, C., BLANK, I. & MÄRK, T. D. 2005. Unambiguous identification of volatile organic compounds by proton-transfer reaction mass spectrometry coupled with GC/MS. *Anal. Chem.,* 77**,** 4117-4124.

LINDINGER, W., HANSEL, A. & JORDAN, A. 1998. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Processes,* 173**,** 191 – 241.

LINSTROM, P. J. & MALLARD, W. G. December 2013. NIST Chemistry WebBook. *Number 69.* NIST Standard Reference Database.

LIU, P., JENKINS, N. A. & COPELAND, N. G. 2003. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res. ,* 13**,** 476-484.

LOPEZ, J. M., THOMS, B. & REHBEIN, H. 1975. Acetoin degradation in *Bacillus subtilis* by direct oxidative cleavage. *Eur. J. Biochem.,* 57**,** 425-430.

LOURENÇO, C. & TURNER, C. 2014. Breath analysis in disease diagnosis: methodological considerations and applications. *Metabolites,* 4**,** 465-498.

MANOLIS, A. 1983. The diagnostic potential of breath analysis. *Clin. Chem. ,* 29**,** 5-15.

MARCHESI, J. R., DUTILH, B. E., HALL, N., PETERS, W. H. M., ROELOFS, R., BOLEIJ, A. & TJALSMA, H. 2011. Towards the human colorectal cancer microbiome. *PLOS ONE,* 6**,** 1-8.

MARKES 2013/2014. Thermal desorption: accessories & consumables.

MARKES 2014. Application note 007 - Calibration: preparing and introducing thermal desorption standards using sorbent tubes.

MATERIĆ, D., LANZA, M., SULZER, P., HERBIG, J., BRUHN, D., TURNER, C., MASON, N. & GAUCI, V. 2015. Monoterpene separation by coupling proton transfer reaction time-of-flight mass spectrometry with fastGC. *Anal. Bioanal. Chem.,* 407**,** 7757-7763.

MCMURRAY, F. & COX, R. D. 2011. Mouse models and type 2 diabetes: translational opportunities. *Mamm. Genome,* 22**,** 390-400.

MCORIST, A. L., ABELL, G. C. J., COOKE, C. & NYLAND, K. 2008. Bacterial population dynamics and faecal short-chain fatty acids (SCFA) concentrations in healthy humans. *Brit. J. Nutr.,* 100**,** 138-146.

MIEKISCH, W., HERBIG, J. & SCHUBERT, J. K. 2012. Data interpretation in breath biomarker research: pitfalls and directions. *J. Breath Res.,* 6**,** 1-10.

MIEKISCH, W. & SCHUBERT, J. K. 2006. From highly sophisticated analytical techniques to life-saving diagnostics: Technical developments in breath analysis. *Trends Anal. Chem,* 25**,** 665-673.

MIEKISCH, W., SCHUBERT, J. K. & NOELDGE-SCHOMBURG, G. 2004. Diagnostic potential of breath analysis—focus on volatile organic compounds. *Clin. Chim. Acta,* 347**,** 25-39.

MINH, T., OLIVER, S., NGO, J., FLORES, R., MIDYETT, R., MEINARDI, S., CARLSON, M. K., ROWLAND, F. S., BLAKE, D. R. & GALASSETTI, P. R. 2011. Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects. *Am. J. Physiol. Endocrinol. Metab.,* 300**,** 1166-1175.

MOCHALSKI, P., KING, J., UNTERKOFLER, K. & AMANN, A. 2013. Stability of selected volatile breath constituents in Tedlar, Kynar and Flexfilm sampling bags. *Analyst,* 138**,** 1405-1418.

MOCHALSKI, P., UNTERKOFLER, K., HINTERHUBER, H. & AMANN, A. 2014a. Monitoring of selected skin-borne volatile markers of entrapped humans by selective reagent ionization time of flight mass spectrometry in NO^+ Mode. *Anal. Chem.,* 86**,** 3915-3923.

MOCHALSKI, P., UNTERKOFLER, K., ŠPANĚL, P., SMITH, D. & AMANN, A. 2014b. Product ion distributions for the reactions of NO^+ with some physiologically significant aldehydes obtained using a SRI-TOF-MS instrument. *Int. J. Mass Spectrom.,* 363**,** 23-31.

MOCHALSKI, P., WZOREK, B., SLIWKA, I. & AMANN, A. 2009. Suitability of different polymer bags for storage of volatile sulphur compounds relevant to breath analysis. *J. Chromatogr. B,* 877**,** 189-196.

MOORE, J. G., JESSOP, L. D. & OSBORNE, D. N. 1987. Gas-chromatographic and mass-spectrometric analysis of the odor of human feces. *Gastroenterology,* 93**,** 1321-1329.

MURPHY, E. F., COTTER, P. D., HEALY, S., MARQUES, T. M., O'SULLIVAN, O., FOUHY, F., CLARKE, S. F., O'TOOLE, P. W., QUIGLEY, E. M., STANTON, C., ROSS, P. R., O'DOHERTY, R. M. & SHANAHAN, F. 2010. Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut,* 59**,** 1635-1642.

MUSA-VELOSO, K., LIKHODII, S. & CUNNANE, S. 2002. Breath acetone is a reliable indicator of ketosis in adults consuming ketogenic meals. *Am. J. Clin. Nutr.,* 76**,** 65-70.

NIV, Y. & SPERBER, A. D. 1995. Sensitivity, specificity, and predictive value of fecal occult blood testing (Hemoccult II) for colorectal neoplasia in symptomatic patients: a prospective study with total colonoscopy. *Am. J. Gastroenterol.,* 90**,** 1974-1977.

NOVEROSKE, J. K., WEBER, J. S. & JUSTICE, M. J. 2000. The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mamm. Genome,* 11**,** 478-483.

NYANGALE, E. P., MOTTRAM, D. S. & GIBSON, G. R. 2012. Gut microbial activity, implications for health and disease: the potential role of metabolite analysis. *Journal of Proteome Research,* 11**,** 5573-5585.

O'KEEFE, S. J., CHUNG, D., MAHMOUD, N., SEPULVEDA, A. R., MANAFE, M., ARCH, J., ADADA, H. & VAN DER MERWE, T. 2007. Why do African Americans get more colon cancer than native Africans? *J. Nutr. ,* 137**,** 175S-182S.

O'HARA, M. E., CLUTTON-BROCK, T. H., GREEN, S. & MAYHEW, C. A. 2009. Endogenous volatile organic compounds in breath and blood of healthy volunteers: examining breath analysis as a surrogate for blood measurements. *J. Breath Res.,* 3**,** 1-10.

O'HARA, M. E., O'HEHIR, S., GREEN, S. & MAYHEW, C. A. 2008. Development of a protocol to measure volatile organic compounds in human breath: a comparison of rebreathing and on-line single exhalations using proton transfer reaction mass spectrometry. *Physiol. Meas.,* 29**,** 309-330.

O'MAHONY, S. M., CLARKE, G., BORRE, Y. E., DINAN, T. G. & CRYAN, J. F. 2015. Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav. Brain Res.,* 277**,** 32-48.

OHIGASHI, S., SUDO, K., KOBAYASHI, D., TAKAHASHI, O., TAKAHASHI, T., ASAHARA, T., NOMOTO, K. & ONODERA, H. 2013. Changes of the intestinal microbiota, short chain fatty acids, and fecal pH in patients with colorectal cancer. *Dig. Dis. Sci.,* 58**,** 1717-1726.

PALLANT, J. 2013. *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS* Berkshire, Open University Press, McGraw-Hill Education.

PAULING, L., ROBINSON, A. B., TERANISHI, R. & CARY, P. 1971. Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proc. Nat. Acad. Sci. USA,* 68**,** 2374-2376.

PERL, T., CARSTENS, E., HIRN, A., QUINTEL, M., VAUTZ, W., NOLTE, J. & JÜNGER, J. 2009. Determination of serum propofol concentrations by breath analysis using ion mobility spectrometry. *Brit. J. Anaesth.,* 103**,** 822-827.

PERRY, R. J., PENG, L., BARRY, N. A., CLINE, G. W., ZHANG, D., CARDONE, R. L., PETERSEN, K. F., KIBBEY, R. G., GOODMAN, A. L. & SHULMAN, G. I. 2016. Acetate mediates a microbiome–brain–β-cell axis to promote metabolic syndrome. *Nature,* 534**,** 213-217.

PHILLIPS, M., CATANEO, R. N., GREENBERG, J., MUNAWAR, M. I., NACHNANI, S. & SAMTANI, S. 2005. Pilot study of a breath test for volatile organic compounds associated with oral malodor: evidence for the role of oxidative stress. *Oral dis.,* 11**,** 32-34.

PHILLIPS, M., GREENBERG, J. & SABAS, M. 1994. Alveolar gradient of pentane in normal human breath *Free Radical Res.,* 20**,** 333-337.

PLEIL, J. & LINDSTROM, A. B. 1998. Sample timing and mathematical considerations for modeling breath elimination of volatile organic compounds. *Risk Anal.,* 18**,** 585-602.

PLEIL, J. D. & LINDSTROM, A. B. 1995. Collection of a single alveolar exhaled breath for volatile organic-compounds analysis *Am. J. Ind. Med.,* 28**,** 109-121.

POOLE, C. F. 2012. *Gas chromatography,* Oxford, Elsevier.

POSCIA, A., MASCINI, M., MOSCONE, D., LUZZANA, M., CARAMENTI, G., CREMONESI, P., VALGIMIGLI, F., BONGIOVANNI, C. & VARALLI, M. 2003. A microdialysis technique for continuous subcutaneous glucose monitoring in diabetic patients (part 1). *Biosens. Bioelectron.,* 18**,** 891-898.

PRAZELLER, P., KARL, T., JORDAN, A., HOLZINGER, R., HANSEL, A. & LINDINGER, W. 1998. Quantification of passive smoking using proton-transfer-reaction mass spectrometry. *Int. J. Mass Spectrom.,* 178**,** L1-L4.

PROUDMAN, C. J., HUNTER, J. O., DARBY, A. C., ESCALONA, E. E., BATTY, C. & TURNER, C. 2015. Characterisation of the faecal metabolome and microbiome of thoroughbred racehorses. *Equine Vet. J.,* 47**,** 580-586.

PYE, G., EVANS, D. F., LEDINGHAM, S. & HARDCASTLE, J. D. 1990. Gastrointestinal intraluminal pH in normal subjects and those with colorectal adenoma or carcinoma. *Gut,* 31**,** 1355-1357.

PYSANENKO, A., SPANĚL, P. & SMITH, D. 2008. A study of sulfur-containing compounds in mouth- and nose-exhaled breath and in the oral cavity using selected ion flow tube mass spectrometry. *J. Breath Res.,* 2**,** 1-13.

PYSANENKO, A., WALTON, T., ŠPANĚL, P. & SMITH, D. 2009. Acetone, butanone, pentanone, hexanone and heptanone in the headspace of aqueous solution and urine studied by selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.,* 23**,** 1097-1104.

QIN, J., LI, Y., CAI, Z., LI, S., ZHU, J., ZHANG, F., LIANG, S., ZHANG, W., GUAN, Y., SHEN, D., PENG, Y., ZHANG, D., JIE, Z., WU, W., QIN, Y., XUE, W., LI, J., HAN, L., LU, D., WU, P., DAI, Y., SUN, X., LI, Z., TANG, A., ZHONG, S., LI, X., CHEN, W., XU, R., WANG, M., FENG, Q., GONG, M., YU, J., ZHANG, Y., ZHANG, M., HANSEN, T., SANCHEZ, G., RAES, J., FALONY, G., OKUDA, S., ALMEIDA, M., LECHATELIER, E., RENAULT, P., PONS, S., BATTO, J.-M., ZHANG, Z., CHEN, H., YANG, R., ZHENG, W., LI, S., YANG, H., WANG, J., EHRLICH, S. D., NIELSEN, R., PEDERSON, O., KRISTIANSEN, K. & WANG, J. 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature,* 490**,** 55-60.

RABOT, S., MEMBREZ, M., BRUNEAU, A., GERARD, P., HARACH, T., MOSER, M., RAYMOND, F., MANSOURIAN, R. & CHOU, C. J. 2010. Germ-free C57BL/6J mice are resistant to high-fat-diet-induced insulin resistance and have altered cholesterol metabolism. *FASEB J.,* 24**,** 4948-4959.

RAHAT-ROZENBLOOM, S., FERNANDES, J., GLOOR, G. B. & WOLEVER, T. M. S. 2014. Evidence for greater production of colonic short-chain fatty acids in overweight than lean humans. *Int. J. Obes.,* 38**,** 1525-1531.

REICHARD, G. A. J., SKUTCHES, C. L., HOELDTKE, R. D. & OWEN, E. 1986. Acetone metabolism in humans during diabetic ketoacidosis. *Diabetes,* 35**,** 668-674.

RIGHETTONI, M., SCHMID, A., AMANN, A. & PRATSINIS, S. E. 2013. Correlations between blood glucose and breath components from portable gas sensors and PTR-TOF-MS. *J. Breath Res.,* 7**,** 1-9.

ROMANO, A., FISCHER, L., HERBIG, J., CAMPBELL-SILLS, H., COULON, J., LUCAS, P., CAPPELLIN, L. & BIASIOLI, F. 2014. Wine analysis by fastGC proton-transfer reaction-time-of-flight-mass spectrometry. *Int. J. Mass Spectrom.,* 369**,** 81-86.

RUZSANYI, V., MOCHALSKI, P., SCHMID, A., WIESENHOFER, H., KLIEBER, M., HINTERHUBER, H. & AMANN, A. 2012. Ion mobility spectrometry for detection of skin volatiles. *J. Chromatogr. B,* 911**,** 84-92.

SALERNO-KENNEDY, R. & CASHMAN, K. D. 2005. Potential applications of breath isoprene as a biomarker in modern medicine: a concise overview. *Wien. Klin. Wochenschr.,* 117**,** 180-186.

SALINGER, A. P. & JUSTICE, M. J. 2008. Mouse mutagenesis using *N*-Ethyl-*N*-Nitrosourea (ENU). *CSH Protocols.* doi: 10.1101/pdb.prot4985.

SÁNCHEZ-LÓPEZ, J. A., ZIMMERMANN, R. & YERETZIAN, C. 2014. Insight into the time-resolved extraction of aroma compounds during espresso coffee preparation: online monitoring by PTR-ToF-MS. *Analytical Chemistry,* 86**,** 11696-11704.

SANDBERG, A., SKOLD, C. M., GRUNEWALD, J., EKLUND, A. & WHEELOCK, A. M. 2011. Assessing recent smoking status by measuring exhaled carbon monoxide levels *PLOS ONE,* 6**,** 1-7.

SANDER, R. 1999. *Compilation of Henry's Law constants for inorganic and organic species of potential importance in environmental chemistry (Version 3)* [Online]. http://www.henrys-law.org/. [Accessed].

SCHUBERT, J., SPITTLER, K. H., BRAUN, G., GEIGER, K. & GUTTMANN, J. 2001. CO_2 - controlled sampling of alveolar gas in mechanically ventilated patients *J. Appl. Physiol.,* 90**,** 486-492.

SCHUBERT, J. K., MIEKISCH, W., BIRKEN, T., GEIGER, K. & NOELDGE-SCHOMBURG, G. 2005. Impact of inspired substance concentrations on the results of breath analysis in mechanically ventilated patients. *Biomarkers,* 10**,** 138-152.

SCHUBERT, R., SCHWOEBEL, H., MAU-MOELLER, A., BEHRENS, M., FUCHS, P., SKLORZ, M., SCHUBERT, J. K., BRUHN, S. & MIEKISCH, W. 2012. Metabolic monitoring and assessment of anaerobic threshold by means of breath biomarkers. *Metabolomics,* 8**,** 1069-1080.

SCHWARZ, K., PIZZINI, A., ARENDACK, B., ZERLAUTH, K., FILIPIAK, W., SCHMID, A., DZIEN, A., NEUNER, S., LECHLEITNER, M., S SCHOLL-B¨URGI7, MIEKISCH, W., SCHUBERT, J., UNTERKOFLER, K., WITKOVSK´Y, V., GASTL, G. & AMANN, A. 2009. Breath acetone—aspects of normal physiology related to age and gender as determined in a PTR-MS study. *J. Breath Res.,* 3**,** 1-9.

SEITZ, H. K., SIMANOWSKI, U. A., GARZON, F. T., RIDEOUT, J. M., PETERS, T. J., KOCH, A., BERGER, M. R., EINECKE, H. & MAIWALD, M. 1990. Possible role of acetaldehyde in ethanol-related rectal cocarcinogenesis in the rat. *Gastroenterology,* 98**,** 406-413.

SHEN, X. J., RAWLS, J. F., RANDALL, T., BURCAL, L., MPANDE, C. N., JENKINS, N., JOVOV, B., ABDO, Z., SANDLER, R. S. & KEKU, T. O. 2010. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut microbes,* 1(3)**,** 138-147.

SHICHIRI, M., KAWAMORI, R., YAMASAKI, Y., HAKUI, N. & ABE, H. 1982. Wearable artificial endocrine pancreas with needle-type glucose sensor. *Lancet,* 2**,** 1129-1131.

SHIN, N. R., LEE, J. C., LEE, H. Y., KIM, M. S., WHON, T. W., LEE, M. S. & BAE, J. W. 2014. An increase in the *Akkermansia spp.* population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut,* 63**,** 727-735.

SHURMER, H., FARD, A., BARKER, J., BARTLETT, P., DODD, G. & HAYAT, U. 1987. Development of an electronic nose *Phys. Technol.,* 18**,** 170-176.

SIMMONS, J., MCFANN, K., BROWN, A., REWERS, A., FOLLANSBEE, D., TEMPLE-TRUJILLO, R. & KLINGENSMITH, G. 2007. Reliability of the diabetes fear of injecting and self-testing questionnaire in pediatric patients with type 1 diabetes. *Diabetes Care,* 30**,** 987-988.

SMITH, D. 1992 The ion chemistry of interstellar clouds. *Chem. Rev.,* 92**,** 1473 – 1485.

SMITH, D. & ADAMS, N. G. 1987. The Selected Ion Flow Tube (SIFT): Studies of ion-neutral reactions. *Adv. Atom. Mol. Phys,* 24**,** 1 – 49.

SMITH, D., CHIPPENDALE, T. W. E., DRYAHINA, K. & ŠPANĚL, P. 2013. SIFT-MS analysis of nose-exhaled breath; mouth contamination and the influence of exercise. *Curr. Anal. Chem.,* 9**,** 565-575.

SMITH, D., ISMAIL, K., DISKIN, A. M., CHAPMAN, G., MAGNAY, J., ŠPANĚL, P. & O'BRIEN, S. 2006. Increase of acetone emitted by urine in relation to ovulation *Acta Obstet. Gynecol.,* 85**,** 1008-1011.

SMITH, D., PYSANENKO, A. & ŠPANĚL, P. 2010a. Kinetics of ethanol decay in mouth- and nose-exhaled breath measured on-line by selected ion flow tube mass spectrometry following varying doses of alcohol. *Rapid Commun. Mass Spectrom.,* 24**,** 1066-1074.

SMITH, D. & ŠPANĚL, P. 1996a. Application of ion chemistry and the SIFT technique to the quantitative analysis of trace gases in air and on breath. *Int. Rev. Phys. Chem.,* 15**,** 231–271.

SMITH, D. & ŠPANĚL, P. 1996b. The novel selected-ion flow tube approach to trace gas analysis of air and breath. *Rapid. Commun. Mass. Spectrom.,* 10**,** 1183-1198.

SMITH, D. & ŠPANĚL, P. 2005. Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass. Spectrom. Rev.,* 24**,** 661-700.

SMITH, D. & ŠPANĚL, P. 2007. The challenge of breath analysis for clinical diagnosis and therapeutic monitoring. *Analyst,* 132**,** 390 – 396.

SMITH, D. & ŠPANĚL, P. 2011. Direct, rapid quantitative analyses of BVOCs using SIFT-MS and PTR-MS obviating sample collection. *Trends Anal. Chem,* 30**,** 945-959.

SMITH, D. & ŠPANĚL, P. 2015. SIFT-MS and FA-MS methods for ambient gas phase analysis: developments and applications in the UK. *Analyst,* 140**,** 2573-2591.

SMITH, D., ŠPANĚL, P., ENDERBY, B., LENNEY, W., TURNER, C. & DAVIES, S. 2010b. Isoprene levels in the exhaled breath of 200 healthy pupils within the age range 7–18 years studied using SIFT-MS. *J. Breath Res.,* 4**,** 1-7.

SMITH, D., ŠPANĚL, P., FRYER, A., HANNA, F. & FERNS, G. 2011. Can volatile compounds in exhaled breath be used to monitor control in diabetes mellitus? *J. Breath Res.,* 5**,** 1-8.

SMITH, D., TURNER, C. & ŠPANĚL, P. 2007. Volatile metabolites in the exhaled breath of healthy volunteers: their levels and distributions. *J. Breath Res.,* 1**,** 1-12.

SMITH, D., WANG, T., PYSANENKO, A. & ŠPANĚL, P. 2008. A selected ion flow tube mass spectrometry study of ammonia in mouth- and nose-exhaled breath and in the oral cavity. *Rapid Commun. Mass Spectrom.,* 22**,** 783-789.

SMITH, D., WANG, T. & ŠPANĚL, P. 2002. On-line, simultaneous quantification of ethanol, some metabolites and water vapour in breath following the ingestion of alcohol. *Physiol. Meas.,* 23**,** 477-489.

SMITH, D. Š., P.; HERBIG, J.; BEAUCHAMP, J. 2014. Mass spectrometry for real-time quantitative breath analysis. *J. Breath Res.,* 8**,** 1-10.

SMITH, F. L. & HARVEY, A. H. 2007. Avoid Common pitfalls when using henry's law. *CEP***,** 33-39.

SOCIETY, A. T. 1999. Recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide in adults and children - 1999. *Am. J. Respir. Crit. Care Med.,* 160**,** 2104-2117.

SOCIETY, A. T. 2005. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am. J. Respir. Crit. Care Med.,* 171**,** 912-930.

SONG, S. H. & HARDISTY, C. A. 2008. Early-onset type 2 diabetes mellitus: an increasing phenomenon of elevated cardiovascular risk. *Expert Rev. Cardiovasc. Ther.,* 6**,** 315-322.

ŠPANĚL, P., DRYAHINA, K., REJSKOVA, A., CHIPPENDALE, T. W. E. & SMITH, D. 2011. Breath acetone concentration; biological variability and the influence of diet. *Physiol. Meas.,* 32**,** 23-31.

ŠPANĚL, P., DRYAHINA, K. & SMITH, D. 2007a. Acetone, ammonia and hydrogen cyanide in exhaled breath of several volunteers aged 4–83 years. *J. Breath Res.,* 1**,** 1-4.

ŠPANĚL, P., DRYAHINA, K. & SMITH, D. 2007b. Microwave plasma ion sources for selected ion flow tube mass spectrometry: optimizing their performance and detection limits for trace gas analysis. *Int. J. Mass Spectrom.,* 267**,** 117-124.

ŠPANĚL, P., DRYAHINA, K. & SMITH, D. 2013. A quantitative study of the influence of inhaled compounds on their concentrations in exhaled breath. *J. Breath Res.,* 7**,** 1-10.

ŠPANĚL, P., JI, Y. & SMITH, D. 1997. SIFT studies of the reactions of H_3O^+, NO^+ and O_2^+ with a series of aldehydes and ketones. *Int. J. Mass Spectrom.,* 165/166**,** 25-37.

ŠPANĚL, P., PAVLIK, M. & SMITH, D. 1995. Reactions of H_3O^+ and OH^- ions with some organic molecules; applications to trace gas analysis in air. *Int. J. Mass. Spectrom. Ion Processes,* 145**,** 177-186.

ŠPANĚL, P., ROLFE, P., RAJANT, B. & SMITH, D. 1996. The selected ion flow tube (SIFT) – A novel technique for biological monitoring. *Ann. occup. Hyg.,* 40**,** 615 – 626.

ŠPANĚL, P. & SMITH, D. 1996. Selected ion flow tube: a technique for quantitative trace gas analysis of air and breath. *Med. Biol. Eng.Comput.,* 34**,** 409–19.

ŠPANĚL, P. & SMITH, D. 1997. SIFT studies of the reactions of H_3O^+, NO^+ and O_2^+ with a series of alcohols. *Int. J. Mass Spectrom.,* 167/168**,** 375-388.

ŠPANĚL, P. & SMITH, D. 1998a. Selected ion flow tube studies of the reactions of H_3O^+, NO^+, and O_2^+ with several amines and some other nitrogen-containing molecules *Int. J. Mass Spectrom.,* 176**,** 201-211.

ŠPANĚL, P. & SMITH, D. 1998b. Selected ion flow tube studies of the reactions of H_3O^+, NO^+, and O_2^+ with some organosulphur molecules. *Int. J. Mass Spectrom.,* 176**,** 167-176.

ŠPANĚL, P. & SMITH, D. 1998c. SIFT studies of the reactions of H_3O^+, NO^+ and O_2^+ with a series of volatile carboxylic acids ans esters. *Int. J. Mass Spectrom. Ion Processes,* 172**,** 137-147.

ŠPANĚL, P. & SMITH, D. 1998d. SIFT studies of the reactions of H_3O^+, NO^+ and O_2^+ with several ethers. *Int. J. Mass Spectrom. Ion Processes,* 172**,** 239-247.

ŠPANĚL, P. & SMITH, D. 1999. Selected Ion Flow Tube – Mass Spectrometry: Detection and Real-time Monitoring of Flavours Released by Food Products. *Rapid Commun. Mass Spectrom.,* 13**,** 585-596.

ŠPANĚL, P. & SMITH, D. 2000a. Influence of water vapour on selected ion flow tube mass spectrometric analyses of trace gases in humid air and breath. *Rapid Commun. Mass Spectrom.,* 14**,** 1898-1906.

ŠPANĚL, P. & SMITH, D. 2000b. Quantification of hydrogen sulphide in humid air by selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.,* 14**,** 1136-1140.

ŠPANĚL, P. & SMITH, D. 2007. Selected ion flow tube mass spectrometry for on-line trace gas analysis in biology and medicine. *Eur. J. Mass. Spectrom.,* 13**,** 77 – 82.

ŠPANĚL, P. & SMITH, D. 2009. Influence of weakly bound adduct ions on breath trace gas analysis by selected ion flow tube mass spectrometry (SIFT-MS). *Int. J. Mass Spectrom.,* 280**,** 128-135.

ŠPANĚL, P. & SMITH, D. 2013. Advances in On-line Absolute Trace Gas Analysis by SIFT-MS. *Curr. Anal. Chem.,* 9**,** 525-539.

ŠPANĚL, P., TURNER, C., WANG, T., BLOOR, R. & SMITH, D. 2006. Generation of volatile compounds on mouth exposure to urea and sucrose: implications for exhaled breath analysis. *Physiol. Meas.,* 27**,** N7-N17.

ŠPANĚL, P., WANG, T. & SMITH, D. 2002. A selected ion flow tube, SIFT, study of the reactions of H3O+, NO+ and O2+ ions with a series of diols. *Int. J. Mass Spectrom.,* 218**,** 227-236.

SPARKMAN, O. D., PENTON, Z. & G., K. F. 2011. *Gas chromatography and mass spectrometry: a practical guide* Oxford, Elsevier.

SPESYVYI, A., SMITH, D. & ŠPANĚL, P. 2015. Selected ion flow-drift tube mass spectrometry: quantification of volatile compounds in air and breath. *Anal. Chem.,* 87**,** 12151-12160.

SRINIVASAN, K. & RAMARAO, P. 2007. Animal models in type 2 diabetes research: An overview. *Indian J. Med. Res.,* 125**,** 451-472.

SRINIVASAN, V., PAMULA, V. K., POLLACK, M. G. & FAIR, R. B. 2003. Clinical diagnostics on human whole blood, plasma, serum, urine, saliva, sweat, and tears on a digital microfluidic platform. *7th International Conference on Miniaturized Chemical and Biochemical Analysis Systems.*

STEERENBERG, P. A., NIERKENS, S., VAN LOVEREN, H. & VAN AMSTERDAM, J. G. C. 2000. A simple method to sample exhaled NO not contaminated by ambient NO from children and adults in epidemiological studies. *Nitric Oxide,* 4**,** 168-174.

STORER, M., DUMMER, M., LUNT, H., SCOTTER, J., MCCARTIN, F., COOK, J., SWANNEY, M., KENDALL, D., LOGAN, F. & EPTON, M. 2011. Measurement of breath acetone concentrations by selected ion flow tube mass spectrometry in type 2 diabetes. *J. Breath Res.,* 5**,** 1-5.

SU, T. 1994. Parametrization of kinetic energy dependences of ion-polar molecule collision rate constants by trajectory calculations. *J. Chem. Phys.,* 100**,** 4703.

SU, T. & BOWERS, M. T. 1973. Ion-polar molecule collisions: the effect of ion size on ion-polar molecule rate constants; the parameterization of the average-dipole-orientation theory. *Int. J. Mass Spectrom. Ion Phys.,* 12**,** 347-356.

SU, T. & CHESNAVICH, W. 1982. Parametrization of the ion-polar molecule collision rate constant by trajectory calculations. *J. Chem. Phys.,* 76**,** 5183-5185.

SULZER, P., HARTUNGEN, E., HANEL, G., FEIL, S., WINKLER, K., MUTSCHLECHNER, P., HAIDACHER, S., SCHOTTKOWSKY, R., GUNSCH, D., SEEHAUSER, H., STRIEDNIG, M., JÜRSCHIK, S., BREIEV, K., LANZA, M., HERBIG, J., MÄRK, L., MÄRK, T. D. & JORDAN, A. 2014. A proton transfer reaction-quadrupole interface time-of-flight mass spectrometer (PTR-QiTOF): high speed due to extreme sensitivity. *Int. J. Mass Spectrom.,* 368**,** 1-5.

SULZER, P., PETERSSON, F., AGARWAL, B., BECKER, K. H., JÜRSCHIK, S., MÄRK, T. D., PERRY, D., WATTS, P. & MAYHEW, C. A. 2012. Proton Transfer Reaction Mass Spectrometry and the Unambiguous Real-Time Detection of 2,4,6 Trinitrotoluene. *Anal. Chem.,* 84**,** 4161-4166.

SZYMCZAK, W., ROZMAN, J., HÖLLRIEGL, V., KISTLER, M., KELLER, S., PETERS, D., KNEIPP, M., SCHULZ, H., HOESCHEN, C., KLINGENSPOR, M. & ANGELIS, M. H. 2014. Online breath gas analysis in unrestrained mice by hs-PTR-MS. *Mamm. Genome,* 25**,** 129-140.

TAKAHASI, K. R., SAKURABA, Y. & GONDO, Y. 2007. Mutational pattern and frequency of induced nucleotide changes in mouse ENU mutagenesis. *BMC Mol. Biol.,* 8:52**,** 1-10.

TANI, A., HAYWARD, S. & HEWITT, C. N. 2003. Measurement of monoterpenes and related compounds by proton transfer reaction-mass spectrometry (PTR-MS). *Int. J. Mass Spectrom.,* 223-224**,** 561-578.

TAUCHER, J., HANSEL, A., JORDAN, A., FALL, R., FUTRELL, J. H. & LINDINGER, W. 1997. Detection of isoprene in expired air from human subjects using proton-transfer-reaction mass spectrometry. *Rapid. Commun. Mass. Spectrom.,* 11**,** 1230-1234.

TAUCHER, J., HANSEL, A., JORDAN, A. & LINDINGER, W. 1996. Analysis of compounds in human breath after ingestion of garlic using proton-transfer-reaction mass spectrometry. *J. Agric. Food Chem. ,* 44**,** 3778-3782.

TECNIPLAST. Available: http://www.tecniplast.it/uk/product/metabolic-cages.html [Accessed 7th July 2016].

THALHAMER, B., BUCHBERGER, W. & WASER, M. 2011. Identification of thymol phase I metabolites in human urine by headspace sorptive extraction combined with thermal desorption and gas chromatography mass spectrometry. *J. Pharm. Biomed. Anal.,* 56**,** 64-69.

TOWLER, B., IRWIG, L., GLASZIOU, P., KEWENTER, J., WELLER, D. & SILAGY, C. 1998. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, hemoccult. *BMJ,* 317**,** 559-565.

TREFZ, P., RÖSNER, L., HEIN, D., SCHUBERT, J. K. & MIEKISCH, W. 2013. Evaluation of needle trap micro-extraction and automatic alveolar sampling for point-of-care breath analysis. *Anal. Bioanal. Chem.,* 405**,** 3105-3115.

TURNER, A. F. P., CHEN, B. & PILETSKY, S. A. 1999. In vitro diagnostics in diabetes: meeting the challenge. *Clin. Chem.,* 45**,** 1596-1601.

TURNER, C., PAREKH, B., WALTON, C., ŠPANĚL, P., SMITH, D. & EVANS, M. 2008. An exploratory comparative study of volatile compounds in exhaled breath and emitted by skin using selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.,* 22**,** 526-532.

TURNER, C., ŠPANĚL, P. & SMITH, D. 2006a. A longitudinal study of ammonia, acetone and propanol in the exhaled breath of 30 subjects using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol. Meas.,* 27**,** 321-337.

TURNER, C., ŠPANĚL, P. & SMITH, D. 2006b. A longitudinal study of breath isoprene in healthy volunteers using selected ion flow tube mass spectrometry (SIFT-MS). *Physiol. Meas.,* 27**,** 13-22.

TURNER, C., WALTON, C., HOASHI, S. & EVANS, M. 2009. Breath acetone concentration decreases with blood glucose concentration in type I diabetes mellitus patients during hypoglycaemic clamps. *J. Breath Res.,* 3**,** 1-6.

UEDA, H., IKEGAMI, H., KAWAGUCHI, Y., FUJISAWA, T., NOJIMA, K., BABAYA, N., YAMADA, K., SHIBATA, M., YAMATO, E. & OGIHARA, T. 2000. Age-dependent changes in phenotypes and candidate gene analysis in a polygenic animal model of type II diabetes mellitus: NSY mouse. *Diabetologia,* 43**,** 932-938.

ULANOWSKA, A., KOWALKOWSKI, T., HRYNKIEWICZ, K., JACKOWSKIC, M. & BUSZEWSKIA, B. 2011. Determination of volatile organic compounds in human breath for *Helicobacter pylori* detection by SPME-GC/MS. *Biomed. Chromatogr.,* 25**,** 391-397.

VAITTINEN, O., SCHMIDT, F. M., METSALA, M. & HALONEN, L. 2013. Exhaled Breath Biomonitoring Using Laser Spectroscopy. *Curr. Anal. Chem.,* 9**,** 463-475.

VAN HARREVELD, A. 2003. Odor concentration decay and stability in gas sampling bags. *J. Air & Waste Manage. Assoc.,* 53**,** 51-60.

VAN MUNSTER, I. P. & NAGENGAST, F. M. 1993. The role of carbohydrate fermentation in colon-cancer prevention. *Scand J Gastroenterol Suppl. ,* 28**,** 80-86.

VAUTZ, W., NOLTE, J., BUFE, A., BAUMBACH, J. I. & PETERS, M. 2010. Analyses of mouse breath with ion mobility spectrometry: a feasibility study. *J. Appl. Physiol.,* 108**,** 697-704.

VIGGIANO, A. A. & MORRIS, R. A. 1996. Rotational and vibrational energy effects on ion-molecule reactivity as studied by the VT-SIFDT technique. *J. Phys. Chem.,* 100**,** 19227-19240.

VIGGIANO, A. A., MORRIS, R. A., DALE, F. & PAULSON, J. F. 1990. Kinetic energy, temperature, and derived rotational temperature dependences for the reactions of $Kr^+(^2P_{3/2})$ and Ar^+ with HCl. *J. Chem. Phys.,* 93**,** 1149-1157.

VIGGIANO, A. A., MORRIS, R. A. & PAULSON, J. F. 1988. Rotational temperature dependences of gas phase ion-molecule reactions. *J. Chem. Phys.,* 89**,** 4848-4852.

WALKER, H. K., HALL, W. D. & HURST, J. W. 1990. *Clinical methods: the history, physical, and laboratory examinations,* Boston, Butterworth Publishers.

WANG, A., CASADEI, F., JOHANSEN, A., BUKMAN, H. & EDELENBOS, M. 2016. Emission of volatile organic compounds from healthy and diseased onions. *Acta Hortic,* 1144**,** 333-340.

WANG, T., PYSANENKO, A., DRYAHINA, K. & ŠPANĚL, P. 2008. Analysis of breath, exhaled via the mouth and nose, and the air in the oral cavity. *J. Breath Res.,* 2**,** 1-13.

WANNIER, G. H. 1953. Motion of gaseous ions in strong electric fields. *Bell Syst. Tech. J.,* 32**,** 170-254.

WARNEKE, C., KUCZYNSKI, J., HANSEL, A., JORDAN, A., VOGEL, W. & LINDINGER, W. 1996. Proton transfer reaction mass spectrometry (PTR-MS): propanol in human breath. *Int. J. Mass Spectrom. Ion Processes,* 154**,** 61-70.

WELLS, D., LAWSON, S. & SIRIWARDENA, A. N. 2008. Canine responses to hypoglycemia in patients with type 1 diabetes. *J. Altern. Complement. Med. ,* 14**,** 1235-1241.

WESTHOFF, M., LITTERST, P., FREITAG, L., URFER, W., BADER, S. & BAUMBACH, J.-L. 2009. Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: results of a pilot study. *Thorax,* 64**,** 744-748.

WINDEY, K., PRETER, V. & VERBEKE, K. 2012. Relevance of protein fermentation to gut health. *Mol. Nutr. Food Res.,* 56**,** 184-196.

WINKLER, K., HERBIG, J. & KOHL, I. 2013. Real-time metabolic monitoring with proton transfer reaction mass spectrometry. *J. Breath Res.,* 7**,** 1-8.

XIAO, Z. & XU, P. 2007. Acetoin metabolism in bacteria. *Crit. Rev. Microbiol.,* 33**,** 127-140.

YOO, E. & LEE, S. 2010. Glucose biosensors: an overview of use in clinical practice. *Sensors,* 10**,** 4558-4576.

Appendix A

ELISA assay

Test Name: INSULIN MILLIPORE

ID2: Cat #: EZRMI-13K

ID3: Lot #: 2430979

Basic settings

Measurement type: Absorbance

Microplate name: NUNC 96

Optic settings

No.	Excitation	Emission
1	450	Empty
2	590	Empty

General settings: Bottom optic used

Positioning delay [s]: 0.5

Reading direction: bidirectional, horizontal left to right, top to bottom

Target temperature [°C]: 25

Absorbance (values are displayed as OD)

Y = Bottom + (Top-Bottom) / (1+(EC50/x)^Slope)
Wavelength:

Top 4.354407201

Slope 1.574396683

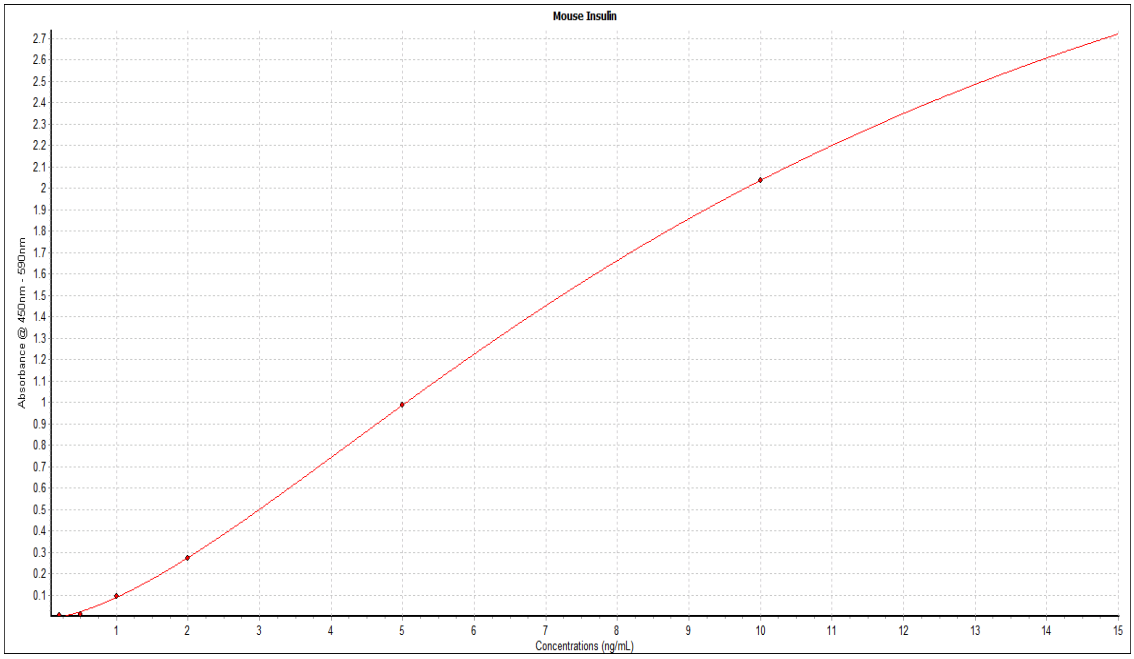
EC50 10.81707518

log(EC50) 1.034109848

Bottom -0.011845421

r = 0.999963023

r^2 = 0.999926047



Appendix B

Metabolites **2014**, *4*, 465-498; doi:10.3390/metabo4020465

OPEN ACCESS

metabolites

ISSN 2218-1989

www.mdpi.com/journal/metabolites/

Review

Breath Analysis in Disease Diagnosis: Methodological Considerations and Applications

Célia Lourenço and Claire Turner *

Department of Life, Health & Chemical Sciences, Chemistry and Analytical Sciences, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK; E-Mail: celia.lourenco@open.ac.uk

* Author to whom correspondence should be addressed; E-Mail: claire.turner@open.ac.uk; Tel.: +44-190-865-2881; Fax: +44-190-865-4167.

Received: 7 February 2014; in revised form: 2 June 2014 / Accepted: 9 June 2014 / Published: 20 June 2014

Abstract: Breath analysis is a promising field with great potential for non-invasive diagnosis of a number of disease states. Analysis of the concentrations of volatile organic compounds (VOCs) in breath with an acceptable accuracy are assessed by means of using analytical techniques with high sensitivity, accuracy, precision, low response time, and low detection limit, which are desirable characteristics for the detection of VOCs in human breath. "Breath fingerprinting", indicative of a specific clinical status, relies on the use of multivariate statistics methods with powerful in-built algorithms. The need for standardisation of sample collection and analysis is the main issue concerning breath analysis, blocking the introduction of breath tests into clinical practice. This review describes recent scientific developments in basic research and clinical applications, namely issues concerning sampling and biochemistry, highlighting the diagnostic potential of breath analysis for disease diagnosis. Several considerations that need to be taken into account in breath analysis are documented here, including the growing need for metabolomics to deal with breath profiles.

Keywords: volatile organic compounds (VOCs); trace gas analysis; volatile biomarkers; metabolomics

1. Introduction

The developments in diagnostic methods and monitoring technologies have focused on blood and urine analysis for clinical diagnostics. The contemporaneous technological advance in analytical

techniques allows the measurement of volatile organic compounds (VOCs) emitted from clinical samples, such as exhaled breath, urine, blood, serum, sputum, and faeces. In spite of its advantages, diagnostics based on VOCs profiling is not yet widely used in clinical practice [1].

During the last decades, Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) [2], Proton Transfer Reaction Mass Spectrometry (PTR-MS) [3], and Gas Chromatography–Mass Spectrometry (GC-MS) [4] with thermal desorption or solid-phase micro extraction (SPME) have been widely used for medical research. SIFT-MS and PTR-MS analytical techniques have been developed for potential medical applications, by using breath analysis, urine analysis, faecal analysis, *in vivo* human skin studies, and *in vitro* cell cultures [5]. SIFT-MS and PTR-MS were developed for real-time, on-line detection and quantification of trace gases in air, with a high sensitivity and wide dynamic range.

Breath is an obvious matrix for analysis of VOCs, as the VOCs are generated within the body, travel around the body via the blood and then they can cross the alveolar interface and appear in exhaled breath, being measured at trace concentrations in the parts-per-million by volume (ppmv) and parts-per-billion by volume (ppbv) levels or lower [6]. Although the trace compounds produced in the oral cavity do not necessarily enter the blood stream, they do appear on exhaled breath. Analysis of the concentrations of VOCs in breath with an acceptable accuracy can provide an indicator of metabolic status, allowing a distinction between healthy and diseased states. Thus, these techniques have the potential to detect diseases in their early stages, non-invasively and painlessly.

Michael Phillips has been a pioneering breath researcher for more than thirty years, providing evidence of the presence of identifiable VOCs in the breath related to lung and breast cancer [4,7,8]. Later on, Anton Amann organized the International Association of Breath Research (IABR), followed by the annual international meetings on Breath Analysis, starting in 2004.

The scientific community is motivated to study all parameters which influence the appearance of VOCs in human breath. For that reason, there are crucial points that should not be neglected, such as standardised methodology for breath sampling and analysis. Factors, such as pulmonary gas exchange and contamination, should be taken into account during the development of breath sampling procedures.

Breath analysis requires elaborate methods of data analysis including multivariate statistical methods, which are applied to show statistically significant differences between the groups (healthy/disease). In addition, there is often little agreement between studies as to which VOCs constitute an appropriate discriminating set. Despite these facts, mass spectrometric analytical techniques have proven to be suited for the challenge and are well suited to both biomarker discovery and whole spectral profiling.

2. Techniques for Breath Analysis

Biomarkers research relies on analytical methods (Table 1) that offer high sensitivity, precision and resolution. The on-line, real-time analytical techniques SIFT-MS and PTR-MS exhibit limit of detection ranging from ppbv to pptv, making them ideally suited to breath analysis [9]. Proton transfer reactions occur in both techniques in a chemical ionization process that allows a very efficient ionization for many organic compounds in the gas phase. Product ion generation in SIFT-MS and PTR-MS is managed using chemical ionization, arising from ion-molecule reactions rather than electron impact or photoionization, with much less fragmentation of the molecules. Thus, these techniques are called *soft* ionisation techniques. SIFT-MS and PTR-MS are well suited to direct, real-time MS

profiling without pre-concentration and with limit of detection ranging from ppbv to pptv. Hence, MS data is well placed for this type of analysis, by means of using sophisticated detectors enabling unequivocal real time quantification of volatile organic compounds, with high sensitivity, precision and resolution. The MS data sets are quite simple, easy to handle, including numerous variables perfectly suited for multivariate statistics. In contrast, a major advantage of chromatographic methods is its very high sensitivity due to sample concentration. In addition, the existence of extensive compounds libraries makes compound identification much easier than in SIFT-MS and PTR-MS.

Table 1. A comparison of the characteristics of the available breath research techniques.

Analytical Method	Mode of operation	Limit of detection (LOD)	Sensitivity	Specificity
SIFT-MS	Direct/Real time	ppbv	High	High
PTR-MS	Direct/Real time	pptv	High	Medium-High
IMS	Real-time	ppbv	Medium	Medium
Sensor arrays	Reference to a database	ppbv	Medium	Medium
GC-MS	Pre-concentration	pptv-ppbv	Very-high	Very-high
LAS	Real-time	ppbv	High	High

Other techniques are also widely used, including laser absorption spectroscopy (LAS), ion mobility spectrometry (IMS), and electronic noses containing a variety of gas sensors and semiconductor-based sensor arrays, although gas sensors are often much less sensitive, usually lack specificity, and are prone to drift. There are also difficulties in inter-device reproducibility. Concerning the IMS technique, the ions are generated by a radioactive strip and they are separated according to their mobilities through the gas, which is usually air at atmospheric pressure [10]. Such devices are not operated at high vacuum conditions and therefore ion-molecule collisions occur, limiting the speed of the ions along the drift tube. Hence, the number of ions reaching the detector is lower compared to the theoretical value. This is an important factor limiting the sensitivity of IMS. During the last ten years it has been applied in medical research, such as detection of skin volatiles [11], detection of volatiles in exhaled breath of patients with lung cancer [12], and determination of anaesthetics concentration in exhaled breath [13]. For the LAS technique, the amount of light absorbed by a sample is related to the concentration of the target specie in the sample. The LAS-based technique cavity ringdown spectroscopy (CRDS) has been successfully applied to measure NO concentration in exhaled breath [14]. This technique enables quantification of volatiles in exhaled breath down to below parts-per-billion by volume levels. It is particularly useful for monitoring purposes, and, recently, the exhaled breath of healthy volunteers was assessed by CRDS [15].

2.1. SIFT-MS

The SIFT technique was conceived and developed by N. G. Adams and D. Smith, in 1976 [16], for the study of ion-neutral reactions at thermal interaction energies [17]. Initially, it was developed to satisfy the need of kinetic data on gas-phase ion-neutral reactions observed in cold interstellar clouds [18]. In 1996, it became a method focused on real-time, on-line analysis of volatile trace gases of biological origin with medical applications [2,19–21], such as clinical diagnosis, therapeutic monitoring, and physiological studies [6,21–23]. With the SIFT–MS technique, it is possible to

identify and differentiate isomers by using three different precursor ions (H_3O^+, NO^+, O_2^+) and applying full-scan mode [24]. Compared to PTR-MS, SIFT-MS is less sensitive due to the existence of a mass filter, which selects the precursor ion to be used according to their mass-to-charge ratio. A clear advantage of SIFT-MS is that no electric field is employed and it is therefore possible to carry out ion-molecule reactions under thermal conditions where the kinetic behaviour is well known. Instead, the precursor ions are produced by electrical discharge, selected by a mass filter according to mass-to-charge ratio, and injected into a fast-flowing carrier gas (Helium), being thermalized [25]. Since the quantification is based on well-understood underlying physics and ion chemistry (in-built kinetics library), there is no need for regular calibration, in contrast to PTR-MS [26].

2.2. PTR-MS

Proton transfer reaction mass spectrometry (PTR-MS) was developed in the mid 1990s by Lindinger and co-workers [27]. This technique allows real-time, on-line determination of absolute concentrations of volatile organic compounds, with detection sensitivity greater than SIFT-MS. The first PTR-MS instruments developed only used H_3O^+ as precursor ion, a disadvantage compared to SIFT-MS. Currently, the latest instruments use a switchable reagent ion capability, alternating between the three precursor ions, H_3O^+, NO^+, and O_2^+ like SIFT-MS [28]. There are usually overlapping ions in clinical sample headspace and the use of PTR–MS equipped with a time-of-flight (TOF) mass analyzer improves mass resolution to assist ion identification [29,30].

PTR-MS employs an electric field, E, along the flow tube axis to increase the velocities of the ions. The change of the ratio E/N, where N is the number density of the drift tube buffer gas molecules, will affect the reagent ion hydration and product ion fragmentation. Under normal operating conditions E/N is in the range 120–130 Td, representing a compromise between reagent ion hydration on the one hand, and molecular (product) ion fragmentation on the other. The electric field also prevents the formation of substantial quantities of cluster ions. In contrast to SIFT-MS, PTR-MS operates at higher effective temperatures and the underlying ion chemistry is often not known [31,32].

The instrument is much shorter, due to the existence of a shorter drift tube, has a typical length 10–20 cm, and, consequently, the pumping system is reduced in size, making PTR-MS suitable for transport. However, the latest advances in SIFT-MS have surpassed this issue, in which the instrument has become much smaller and suitable for transport [26].

The recent advances in PTR-MS technology have demonstrated a diverse range of applications, especially for breath gas analysis [33].

2.3. Electronic Noses and Semiconductor-Based Sensor Arrays

In order to measure different VOCs, many applications have combined various sensors and materials into a single array, leading to the development of an "electronic nose" [34]. Sensor technology has been used for many years in clinical testing. Currently, the goal consists of finding materials with high sensitivity and good selectivity to the VOCs to be detected. Up to now, the existing materials are mostly conductive polymers, semiconducting metal oxides, or a combination of the two [35]. Unique sensors, based on nanoparticles appear as a reliable alternative tool for breath analysis, proving to be inexpensive and easy-to-use. The quartz crystal microbalance (QCM) and the surface acoustic wave

(SAW) device are mass-sensitive sensors which have been used in breath analysis [36]. In QCM, gas molecules are adsorbed on the crystal's surface during sensor exposure to a gaseous medium, changing its mass and resonant frequency. A selective adsorption of the gas mixture strongly influences the degree of crystalline order and by the nanostructure boundaries. The polymer overall characteristics, length, planarity of the conjugate polymer chain, and the side chain composition may influence the polymer conductivity [35].

Recently, Cr- or Si-doped WO_3 nanoparticles have shown high sensitivity to acetone, leading to the development of a portable chemo-resistance sensor suitable for real-time breath acetone detection [37]. Breath acetone concentration of five test persons at rest or during physical activity was measured and compared to that measured by PTR-MS. Si-WO_3 sensors were selective to acetone in realistic conditions (90% Relative Humidity), and able to detect differences in breath acetone concentrations between 880 to 980 ppb, in agreement with PTR-MS measurements.

However, the combination of specificity, selectivity, robustness in operation, reproducible manufacturing uniformity, and long-life stability is not offered by current sensors at an acceptable cost level. These technologies are unable to identify individual compounds, although they can be used to compare samples to see whether they have similar VOC profiles.

3. The Challenge behind the Method

Volatile compounds in breath are produced by metabolic processes at various organs and places in the body, in the oral cavity by bacterial infections, by bacteria in the gut or both. However, is likely that many are not biochemically produced in the body and predominantly come from environmental exposure.

Breath analysis started in the 1970s when Linus Pauling and co-workers detected over 200 different VOCs in human exhaled air and in urine headspace by gas chromatography [38]. Apart from the major components of breath, such as acetone, isoprene, *etc*., many trace compounds present are not of endogenous origin [39,40]. A large variety of trace gases exist in ambient air and these can be taken up via inhalation and skin absorption; the source of others is through ingestion. Hence, some of the trace compounds in exhaled breath, perhaps the majority of them, will be exogenous, and these need to be distinguished from the truly endogenous compounds [40,41]. The endogenous compounds found in human breath, such as inorganic gases (e.g., NO and CO), and VOCs (e.g., isoprene, ethane, pentane, acetone) can be measured directly; other typically non-volatile substances, such as isoprostanes, peroxynitrite, or cytokines, can be measured in breath condensate [39]. These non-volatile substances are supposed to be present in exhaled breath as aerosol particles.

Background air VOC concentrations are an issue. The concept of alveolar gradient proposed by Michael Phillips [42], defined as *the abundance in breath minus the abundance in room air*, for substances having higher inspired than expired concentrations has been proposed to deal with this, however it, does not properly account for the background air a subject breathes. This has been demonstrated by the work of many researchers, including Schubert [43] and Španěl [44]. Indeed, it has been shown [43] that this does not lead to quantitative results; instead they suggested that when inhaled (ambient) concentrations of compounds are greater than 5% of the exhaled concentrations, exhaled concentrations cannot be correlated with blood levels with confidence. The levels detected in

breath will depend on many factors, including the concentration in ambient air, the duration of exposure, the solubility and partition co-efficient into tissues, the mass and fat content of the individual, as well as the underlying endogenous concentration. Sample procedures for breath analysis have many advantages over traditional blood analysis of compounds suspended in or dissolved in blood: they are painless and non-invasive, easy to perform, inexpensive, and the results are available immediately for therapeutic assessments [36]. The exhaled air matrix is less complex than that of blood or other body fluids [41]. However, storage of blood is generally easier than breath. Traditional blood analysis typically involves measuring the concentrations of specific salts, proteins or other non-volatile components. However, sampling procedures for blood analysis are stressful for patients and a non-invasive sample, such as breath, is, therefore, preferable.

3.1. Mouth- vs. Nose-Exhaled Breath

The challenge of breath analysis for clinical diagnosis and therapeutic monitoring lies in the identification of endogenous volatile compounds present in mouth-exhaled breath which are potential markers of diseases, and which reflect levels in the systemic circulation. Many of these trace volatile compounds may be produced in the airways, the oral cavity by bacterial infections, by bacteria in the gut, and also emitted from mucus, saliva and aerosols created in the respiratory tract. Phillips and co-workers performed a pilot study by GC–MS for detecting VOCs in breath associated with oral malodour [45]. However, this technique has some limitations since the lower molecular weight VOCs may not be detected due to the sorbent trap selectivity for two carbon atoms or more.

A sampling device of breath exhaled via the mouth, nose, and air in the mouth cavity was developed by Smith and co-workers [46]. Studies were performed using SIFT–MS to evaluate the concentration of mouth- and nose-exhaled breath, in order to understand the biological origin of several VOCs [46–48]. Ammonia in the exhaled breath is largely generated in the mouth [47] as is ethanol [46] and hydrogen cyanide [46]. This has been demonstrated through showing that the levels measured in the nose exhalations are much lower than those observed in the mouth exhalations. Very low concentrations of propanol and acetaldehyde in exhaled breath appear to be partially systemic and partially mouth generated [46]. Acetone, methanol, and isoprene showed similar profiles for mouth- or nose-exhaled breath [46], indicating that these compounds are totally systemic. However, methanol may be ingested by food intake or drink; hence methanol concentration in breath may not be totally produced through the human biochemistry. Thus to avoid the possibility of contamination of the endogenous VOCs through mouth flora, taking breath from the nose is desirable.

A recent study using solid-phase micro-extraction of bacterial cultures demonstrated that several compounds detected in mouth-exhaled breath are produced by anaerobic bacteria in tongue biofilms [49]. In addition, poor oral hygiene can be a confounding factor leading to production of ammonia from urea or ethanol from sugars, thus, increasing VOCs concentration in mouth-exhaled breath [50].

Furthermore, volatile compounds may be produced by bacteria in the gut, transported to and excreted by the lungs [51]. *Helicobacter pylori* living in the human stomach release VOCs that can be detected in mouth-exhaled air [52].

The analysis of mouth- and nose-exhaled breath following ingestion of different doses of alcohol [53] at different concentrations in water was carried out using SIFT-MS by Smith and co-workers. They

determined how the volume of ingested liquid influenced the gastric retention and degradation of ethanol. This showed the fraction of ethanol ingested and, consequently, the fraction which enters the blood stream, which in turn is diluted in exhaled breath. The decay of breath ethanol has been followed and measured in mouth- and nose-exhaled breath. The authors suggested that saturation of the liver enzymes have an important role in managing the decay of breath ethanol. Additionally, has been shown that gastric retention clearly results in a slower release of ethanol into the gut.

3.2. Physiological Levels of Volatiles

It is well known that human breath is a complex matrix of volatile organic compounds, non-volatile organic compounds (aerosol particles), and inorganic compounds. In order to develop a diagnostic breath test, ready to be used in clinical practice, it is necessary to unravel the baseline physiological levels of volatiles present in human breath, and their relationship with age, gender, ethnicity, and metabolic changes in the body.

3.2.1. Common Breath Metabolites

Exhaled breath consists largely of unmodified (inhaled) nitrogen (somewhat less than 74% by volume if water vapour is included; 79% of the permanent gases) and argon (about 1%), oxygen (reduced from 21% inhaled to about 15% exhaled), carbon dioxide (about 5%), and water vapour (saturated at 37 °C -about 6% [54]).

Along with these major gases and vapours, there are many endogenously formed gaseous volatile metabolites present at trace levels variously measured in parts per million (ppmv), parts per billion (ppbv), and even parts per trillion (pptv).

Diskin and co-workers developed an initial study by SIFT–MS, measuring concentrations of the common breath metabolites ammonia, acetone, isoprene, ethanol, and acetaldehyde in the breath of five healthy subjects over a period of 30 days [40,55]. The mean concentrations were calculated, and meaningful distributions obtained for ammonia, acetone, isoprene, and ethanol.

Later on, Turner and co-workers [56–59] performed longitudinal studies of the common metabolites ammonia, acetone, methanol, ethanol, propanol, acetaldehyde, and isoprene in the breath of 30 healthy volunteers over a six-month period, using SIFT–MS. Thus, the biological variability was assessed and the concentration distributions for these metabolites have been determined on-line in single breath exhalations and showed to be a log normal distribution for these metabolites. Ammonia was shown to be a major breath metabolite with a geometric mean of 833 ppb, followed by acetone (477 ppb), methanol (461 ppb), ethanol (112 ppb), isoprene (106 ppb), propanol (18 ppb), and acetaldehyde (22 ppb) [56–59]. Nevertheless, it has been proved that the majority of ammonia seen in mouth-exhaled breath has its origin in the oral cavity [47]. Ammonia is also produced systemically, it appears in the body as a breakdown product of proteins, a contribution originated from the bacterial degradation of protein in the intestine [57]. The metabolic pathway is originated in the liver, where the ammonia is converted into urea, which is then eliminated in urine. Some of the ammonia is expelled from the breath and some is emitted by the skin [60].

The metabolic pathways of acetone are well established. The decarboxylation of acetoacetate and the dehydrogenation of isopropanol are the two sources of acetone production [61]. Acetone levels are

elevated in diabetes, due to rise of blood sugar level and intensive lipolysis [61]. However, acetone was not reported as a unique biomarker of diabetes.

Methanol and ethanol may arise as anaerobic fermentation products by gut bacteria [62] including all alcohols in the series from methanol to heptanol. Methanol is contained in some foods, such as apples and drinks, which, when ingested, increases the methanol in the circulation and, hence, in the exhaled breath [62]. Methanol is used industrially as a solvent, pesticide, and alternative fuel source. It also occurs naturally in animals and plants. Methanol can be absorbed into the body by inhalation, ingestion, skin contact, or eye contact. Methanol does not appear to be generated in the mouth and levels detected in breath are of systemic origin. Most breath ethanol, however, appears to be due to mouth fermentation of sugars (unless the subject has been consuming alcoholic drinks) [63].

The biochemical origin of isoprene in human breath is not entirely clear. However, it is considered to be a marker of cholesterol synthesis [64]. Abnormal breath isoprene levels are related to end-stage renal failure and increases in isoprene levels have been associated with oxidative stress. However, this assumption has not been proved by the work of Lirk and co-workers [65]. Little focus has been given, thus far, to the relationships between breath levels and the underlying systemic concentrations. For that reason, King and co-workers [66] joined efforts in investigating the potential stores of isoprene in peripheral tissue groups. Their findings suggested that breath isoprene variability during exercise is linked to local variations of gas exchange in peripheral tissues. The observable wash-out behaviour of isoprene was attributed to an increased fractional perfusion of potential storage and production sites.

2-propanol is a product of the enzyme-mediated reduction of acetone. Bacteria present in the gut produce alcohols, including 1-propanol and 2-propanol, structural isomers that exists in the human body [67].

There is increasing evidence that acetaldehyde, rather than alcohol itself, is responsible for the carcinogenic effect of alcohol [68]. The ethanol levels in the exhaled breath are clearly increased after consumption of sugars and the action on it by either mouth or gut flora/enzymes [59,63]. Acetaldehyde levels result from endogenous ethanol metabolism [36]. As a consequence, acetaldehyde concentrations in breath are invariably lower than the corresponding ethanol concentrations. In healthy individuals, it is rapidly cleared by conversion into acetic acid and thus it is present at low concentrations in the body. However, these levels in breath may not be obvious because acetaldehyde can also be produced from cellular activity involving sugars.

3.2.2. Age Influence/Gender

Age and gender of the volunteer revealed to be an important factor to be taken into account in breath analysis. PTR–MS was used for determination of isoprene concentrations in children's exhaled breath by Taucher and co-workers, and recognised to be significantly lower than in adults [69]. Lechner and co-workers measured the VOCs on the breath of 126 volunteers, using the same technique, reporting an increase in isoprene concentration of exhaled air of male subjects [70]. SIFT-MS was also used to determine the concentrations of some metabolites in the breath of healthy children aged between 7–18 years old [67]; where the median concentration of pentanol (15 ppb) was also determined. The exhaled breath of several volunteers within the age range 4–83 years was measured and reported a trend of increasing breath ammonia concentration with age [71]. Isoprene is apparently elevated in

breath during adolescence, as reported by Smith and co-workers, probably due to the onset of puberty, as stated by the authors [72].

3.2.3. Influence of Food

The levels of breath metabolites are influenced by food intake [23,73]. The breath metabolites ammonia, methanol, ethanol, propanol, formaldehyde, acetaldehyde, isoprene, and acetone were quantified by SIFT-MS for a group of five volunteers, before and after ingesting 75 g of glucose in the fasting state [60]. Increased levels in the blood/exhaled breath after the consumption of alcohol were also observed, as reported by Smith and co-workers, using SIFT–MS [63]. Similarly, studies on ethanol metabolism were recently reported by Winkler and co-workers [74]. Metabolic degradation of ethanol was tracked by the ingestion of isotope-labelled ethanol using real-time breath gas analysis with PTR-MS. The findings indicated that in part, ethanol was metabolized to acetone and isoprene, as deuterated acetone and isoprene were observed in the mass spectra. However, the signal of the deuterium-labelled acetaldehyde was not observed, suggesting that this product did not enter the blood stream but was rapidly further metabolized.

The volatile organic compounds emitted by mouth-exhaled breath after garlic ingestion was assessed by PTR-MS, and the main constituents of garlic were reported [75]. The results showed variation in VOCs levels along the time. Other products, such as onion, mint, banana and coffee, are also known to emit volatiles at trace concentrations.

3.2.4. Ovulation

During analysis of the headspace of urine from a number of female volunteers, Smith and co-workers observed acetone and ammonia levels occasionally higher than the normal. The urine samples were collected before any food intake. Such findings suggested that it may be caused by metabolic changes occurring during ovulation and related to menstrual cycle length [76,77].

Studies applied to volatile organic compounds present in exhaled breath correlated with ovulation have not yet been reported.

3.3. Exposure to Volatile Compounds

Smoking and Air Contaminants

It has been shown that the composition of exhaled breath is considerably influenced by exposure to pollution and indoor-air contaminants, for example, smoking-enhanced acetonitrile concentrations were found in the breath and urine of smokers [3,21,78]. Compounds present in cigarette smoke, such as 2,5-dimethylfuran, acetonitrile, benzene, toluene, and styrene, can also be identified in smokers and passive smokers' breath [79,80]. Acetonitrile in breath is a good indicator of whether a given subject is a smoker or not because the concentration of acetonitrile in breath takes nearly a week after cessation of smoking to decrease to that of non-smokers [3]. Such findings were reported in an early study using PTR-MS. Smoking increased exhaled ethane and pentane levels in breath. This may be caused by high concentrations of hydrocarbons in cigarette smoke, as well as oxidative damage caused by smoking [81]. Hydrogen cyanide (HCN), along with acetonitrile and benzene, are known to be present in exhaled

breath [82], after analysing the exhaled breath of smokers compared with non-smokers. Measuring carbon monoxide (CO) in exhaled breath is a well-established method used to differentiate between smokers and non-smokers [83]. As a constituent of cigarette smoke, carbon monoxide enters the blood circulation during smoking and forms carboxyhemoglobin (COHb). The elimination of CO is primarily by respiration, thus, there is a strong correlation between CO in breath and COHb [83].

Aside from smoking, there are other ways of VOCs entering the body. Breath contains a diverse range of VOCs that can be taken up by the body through inhalation or skin, and, depending on distribution kinetics, may be present in exhaled breath for different periods after exposure. For instance, limonene was found in most air fresheners and cleaning products and is emitted by wooden furniture and floorings. It is known to be soluble in blood and adipose tissue and, therefore, has the potential to be taken up by the body during inhalation [84].

Background contaminants [43,44] are an important issue, particularly, when a compound is present in both alveolar breath and inspired air. One approach is to provide a source of purified air [85] to inspire during the breath collection and, in this way, it is possible to determine which of the VOCs have been endogenously originated. However, different compounds have different wash-out periods from the human breath [86], hence, the use of purified air for removal of exogenous compounds is restricted.

4. Sampling and Analysis

The VOCs in breath are at trace levels and that coupled with the high humidity, means that storage and transport of samples is challenging. Therefore, any contamination or improper methodology may have a significant impact on the composition and concentration of VOCs detected in exhaled air. Controlled sampling is a key requirement for reliable analysis of breath biomarkers, because sampling methodology can greatly affect the results. This is a major issue in breath analysis. Now there are no accepted standardized methods for on-line or off-line VOC breath-gas sampling and analysis. The first guidelines concerning sample collection for breath analysis was released in 1999 by the American Thoracic Society for nitric oxide (NO) monitoring in breath [87]. Later on, updated guidelines were published in 2005 for measurement of NO in mouth- or nose-exhaled breath [88]. In 2005, recommendations were also published for exhaled breath condensate sampling and analysis [89]. Hence, reproducibility and reliability of sampling methods and analytical measurement procedures continue to be of critical importance.

Biological variability among subjects has been introduced as an issue in breath sampling. Therefore, breath analysis methods based on monitoring subjects over time may be desirable because they can serve as their own controls [90]. In order to perform breath sampling it is necessary to consider the diffusion of volatile organic compounds from blood to alveolar air, which depends on their physicochemical properties, such as, polarity, solubility in fat, Henry partition constant, and volatility [41].

Temperature dependence strongly influences the composition of breath samples in off-line measurements. As the temperature of breath sample falls below body temperature (37 °C), water vapour condenses on the inside of the bag and takes down water soluble compounds. Warming the sample to body temperature will avoid condensation issues and compound losses due to negative temperature gradients [91]. Further aspects related to sampling continue to be debated in the scientific

community, such as body posture of the subject when providing the breath sample; hyperventilation; control of the flow or volume of breath during collection; sampling via nose or mouth; number of breath samples to be taken to reduce variability (single or multiple breaths); dilution and contamination of the sample; physiological parameters, such as respiratory rate or heart beat rate; alveolar breath or end-tidal volume and dead space; number of subjects per study to avoid over-modelling; and direct analysis or sampling for storage [90].

The number of issues to be considered suggests that the development of several protocols for standardized breath sampling may become mandatory.

4.1. End-Tidal and Alveolar Breath

Concentrations of volatile compounds in blood are reflected by their concentrations in the exhaled air, depending on their blood-gas partition coefficient or solubility. Alveolar breath is the part of exhaled air in equilibrium with systemic blood, whereas end-tidal air is the last fraction of expired air, whose composition resembles alveolar air. Generally, the term end-exhaled breath is applied because it does not imply that the composition of expired air is always identical to the equilibrated air inside the alveoli. It has long been acknowledged that alveolar gas exchange is dependent on ventilation, pulmonary perfusion, and the blood:air partition coefficient [92], thus, the non-homogeneities in the composition of alveolar air among different lung regions over different blood:air solubilities of volatile organic compounds [93]. There is evidence that the gas exchange of highly soluble volatile compounds occurs in the airways rather than alveoli [94], meaning that VOCs measured at the mouth depend on expiratory flow. Generally, the term alveolar breath may be applied for low blood soluble VOCs, whereas for highly soluble volatiles such as acetone the term end-exhaled breath should be used due to the evidence that gas exchange occurs in the airways rather than alveoli. Such evidence was quantified for the first time by Španěl and co-workers, who demonstrated the discrepancy between the concentration in the alveolar region to that in exhaled air (a factor of 3 for isoprene) [44].

Breath analysis for medical diagnosis relies on end-tidal sampling [95], involving the collection of only end-tidal air. Alveolar concentration reflects the concentration in blood and consequently, the concentration in blood reflects the metabolic processes occurring in the body. Furthermore, there is evidence that better reproducibility of data is obtained when only the end-tidal fraction of breath is analysed [96].

Earlier in 1948, it was noted by Fowler that the volume of exhaled air is a mixture of dead space and alveolar air [97]. Dead space was previously defined as the volume of expired air, which acts as a conducting airway (nose, pharynx, larynx, trachea), whereas alveolar air is the expired air fraction that has been exchanged in the alveoli.

Initial approaches have been taken, such as discarding the first 500 mL of exhaled breath to avoid dilution of the sample by dead space. However, it incorrectly assumes that all subjects have the same volume of dead space [98].

Monitoring of expired CO_2 has been used to identify alveolar gas. Thus, Schubert and co-workers measured CO_2 [99,100] in exhaled air of mechanically ventilated patients by means of a capnograph, using a CO_2-triggered alveolar sampling valve. They reported this as a reliable and reproducible method for alveolar sampling through CO_2 concentration measures. Later on, Di Francesco and co-workers designed a CO_2-triggered breath sampler suitable for multiple breaths [96]. More recently,

Filipiak and co-workers [101] applied the same on-line monitoring of expired CO_2 in order to collect alveolar air by needle traps used in GC-MS analysis.

The on-line breath sampling so-called *buffered end-tidal* (BET) [95] breath sampling method has been developed to extend the analysis time of the end-tidal fraction of a single exhalation. This sampling system was designed to buffer only the end-tidal fraction of the breath. The patient is asked to exhale through a tailored tube in which the end-tidal fraction of breath is buffered.

Concentration of breath molecules, prior to mouth appearance, still remains a hard task to unravel due to the lack of knowledge about the parameters and processes which could affect the VOCs' final concentrations in the mouth. For instance, highly soluble compounds are diluted on their way up from the deeper respiratory track to the airway opening, leading to a dilution effect on VOCs' concentrations. Such an important subject is further explained in Section 4.2. The concentrations of breath molecules exhibit flow rate dependency, namely, changes in ventilation strongly influence quantification of volatiles in exhaled breath [102]. In additional, body posture and stress can have a significant impact on the observed breath concentration [102]. Breath holding [93] has been demonstrated to significantly increase the exhaled concentrations of breath gases (H_2, CH_4, and CO).

These sampling systems, which selectively extract end-tidal air by discarding anatomical dead space volume, are far from being perfect, since they do not take into account physiological variability.

4.2. Dilution and Contamination

Hydrophilic exhaled trace gases, such as acetone, interact with the water-like mucus membrane lining the conductive airways, an effect known as wash-in/wash-out behaviour [103]. The exhaled breath concentrations of water soluble substances appear to dilute on their way up from the deeper respiratory track to the airway opening, leading to discrepancies between the true alveolar breath and the measured concentrations, demonstrating a dilution effect. It means that highly soluble gases are present in large concentrations in the airway tissue and mucus as compared to less blood-soluble gases for a given partial pressure. An absorption-desorption phenomenon occurs in the airways; this is firstly initiated by absorption of soluble gases from the airway wall to inspired air, during inspiration. By the time the air reaches the alveoli, the air is saturated with soluble gas and no further gas exchange occurs. During expiration, a gradient air-to-mucus is established promoting the deposition of soluble gas on the mucus and delays the rise in soluble gas partial pressure at the mouth. An anatomic dead space cannot be defined for these gases [92,94].

The airway gas exchange is influenced by perfusion; diffusion through the airway wall; and temperature. Perfusion is driven by the bronchial blood flood, meaning that an increase in blood flow increases the amount of blood soluble gas in the exhaled breath.

Smith and co-workers recently reported a quantitative study, where they investigated the relationship between the exhaled and inhaled air concentrations for seven compounds [44]. The volunteers were deliberately exposed to known concentrations of some compounds, within the range of permissible exposure limits. Their findings were consistent with previous models, and the equilibrium concentration of acetone was shown to be enhanced above the measured exhaled end-tidal value by 19% for all subjects.

4.3. Sampling of Single or Multiple Breaths

Breath sampling may be performed for a single breath or for multiple breath cycles [95,101,104]. However, the composition of a single breath may not be a representative alveolar gas sample for the reason that breaths may considerably vary from each other due to different modes and depth of breathing. Multiple breaths may be preferable in order to acquire reproducible breath samples. However, it may be that exact quantification is less important when characterizing breath profiles. Comparison of rebreathing and on-line single exhalations of highly soluble compounds acetone and methanol, and the low soluble isoprene, was previously evaluated [104,105]. For highly soluble compounds, such as acetone, exchange occurs in the airways rather than alveoli [94], thus, breath sampling becomes much more complicated because acetone concentration in end-exhaled breath may not be in equilibrium with the systemic blood. For that reason, isothermal rebreathing model has been proposed for estimating the alveolar levels of highly soluble exhaled endogenous volatiles [105].

Breathing patterns [102] have been studied and measurements, such as mouth pressure, tidal volume, respiration rate, end-tidal carbon dioxide, and mixed expired carbon dioxide, were recorded. Paced breathing [102] profiles showed reduced breath variability, according to mass and respiration rate. The authors suggested that controlled breathing would prevent hyperventilation, reducing variability in ventilation.

4.4. Storage and Stability of Breath Samples

Direct sampling is preferable to storage for later analysis. This way the decomposition of samples or loss of compounds by diffusion is avoided.

When direct analysis is not possible, the appropriate storage of exhaled breath is an important issue to consider. Background emission of pollutants, losses by diffusion through the bag or adsorption to the inner bag, and interactions between sample constituents, namely reactive chemistry of the stored sample, may irreversibly modify the original sample composition and consequently distort the final results of analyses.

Currently, Tedlar bags are the most common materials for breath collection [106]. Nalophan bags are also popular due to its low price, inertness, and relatively good durability. Generally, breath may be stored in several ways [51]:

- Transparent or black Tedlar bags (PTFE-polytetrafluoroethylene)
- Flexfoil bags (PET/NY/AL/CPE-polyethylene terephthalate/nylon/aluminium foil/chlorinated polyethylene)
- Nalophan bags (PET-polyethylene terephthalate)
- Glass vials (for SPME)
- Thermal desorption tubes (different adsorbents, used in TD GC-MS)
- Micropacked sorbent traps
- Metal canisters

The stability of selected breath constituents in polymer sampling bags have been previously investigated and assessed by PTR-MS and GC-MS [106–108]. Smaller samples are more vulnerable to VOCs losses by permeation. Additionally, the volume of the sample collected affects the stability of

the sample, thus, Mochalski and co-workers recommended sample collections as large as possible to prevent background emissions of contaminants [106]. Previous studies reported the testing of Nalophan bags for storing tobacco samples, and investigated the factors contributing to decay of gas samples during storage, between 4 and 40 h after collection [109]. Samples remained relatively stable between 4 and 12 h after sampling. The odour concentration decreases after 30 h storage to about half of their initial value, due to diffusion effects. Background contaminants released from the bags must be taken into account. Emissions from Nalophan, Flexfoil and Teflon bags were assessed 48 h after filling for storage of volatile sulphur compounds, and none showed emissions of contaminants, thus, all proving to be excellent materials for breath sample storage. Tedlar bags, however, showed significant emissions of COS and CS_2, especially for black Tedlar bags with COS and CS_2 emissions being seen at up to 7 ppb after three days of storage [108]. A study performed by Gilchrist and co-workers investigated the collection and storage of breath samples containing hydrogen cyanide [91]. Breath was collected into 25 μm thick Nalophan, 70 μm Nalophan and Tedlar bags, at 20 °C or 37 °C. Results showed better correlation between on-line and off-line concentrations for all bag types at 37 °C. Correlation of hydrogen cyanide concentrations in breath samples stored at 37 °C was good up to 24 h for the 70 μm Nalophan and Tedlar bags. Such findings suggested that either would be appropriate to use for collection of breath containing hydrogen cyanide. However, Nalophan bags are much cheaper than the Tedlar bags and they can be discarded after a single use, removing the need for bag cleaning or infection control measures.

Humidity also affects the species recoveries, and the high humidity in exhaled breath might cause significant decrease in vapour concentrations for those compounds highly miscible with water [106,107]. Water vapour diffuses through most bags at a speed dependent on the temperature of the bag material [91]. Such findings can easily be tracked by exploiting the full capabilities of SIFT-MS to measure the water vapour in air/breath samples.

Condensation affects the sample authenticity, especially for water-soluble compounds. The loss of volatile compounds to condensed water in Tedlar bags used for breath sampling has been previously evaluated [110] showing differences between dry and wet matrices smaller than 10%. For VOCs with molecular mass above 110 amu, higher losses were detectable (20%–40%) [106]. Thus, Mochalski and co-workers recommended storing breath samples in pre-conditioned Tedlar bags up to 6 h at the maximum possible filling volume.

Recently, needle trap micro-extraction (NTME) [111] combined with GC has been assessed for sample preparation in VOCs analysis. This is a technique similar to SPME but with the advantage of allowing automatic alveolar sampling. The analysis is quite similar to that used in thermal desorption traps. VOCs are thermally desorbed from the needle trap device and separated, identified and quantified by means of two-dimensional gas chromatography combined with MS detector, GC × GC-MS, which has been applied to solve complex problems of separation. Needle traps have offered increased robustness in comparison to SPME, due to the existence of an extraction sorbent packed inside a hypodermic needle rather than supported on a fragile silica fibre that is exposed to the breath matrix during extraction. The influence of humidity, sample volume, and sampling flow has to be thoroughly evaluated in order to be used for pre-concentration of breath volatiles.

4.5. Physiological Parameters

The complex physiological mechanisms underlying pulmonary gas exchange makes breath analysis a challenging subject. Gas exchange [112] during respiration occurs primarily through diffusion. It takes place between the air within the alveoli and the pulmonary capillaries.

Nowadays, it is known that exhalation of breath biomarkers may well depend on physiological parameters, such as blood pressure, heartbeat rate and alveolar ventilation [113]. Exhaled acetone concentrations mirrored exercise induced changes of dextrose metabolism and lipolysis [113]. Understanding the influence of these factors is an essential requisite for the development of a reliable methodology based on breath volatiles. Breath gas concentration can then be related to blood concentrations via mathematical modelling. The simplest model relating breath gas concentration to blood concentrations was developed by Farhi.

Through Farhi [114], Equation (1), is observed that alveolar air concentration, C_A, is proportional to the concentration of VOCs in mixed venous blood (C_V) and depends on: blood:air partition coefficient, $\lambda_{b:air}$, which describes the diffusion equilibrium between capillaries and alveoli, and ventilation-perfusion ratio, $\frac{V_A}{Q_c}$. Such ratio ensures that the ideal amount of blood and gas is received by the alveoli for efficient gas exchange. It depends on the alveolar ventilation (V_A) controlling the transport of the VOC through the respiratory track, and cardiac output (Q_c) controlling the rate at which the VOC is delivered to the lungs. Blood:air partition coefficient, $\lambda_{b:air}$, is strongly dependent on temperature ranging from 23 °C in the mouth to 37 °C in the alveoli, affecting soluble gas exchange [92]. This coefficient represents the ratio of the concentration in blood to the concentration in the gas phase.

$$C_{measured} = C_A = \frac{C_V}{\lambda b{:}\,air + \frac{V_A}{Q_c}} \quad (1)$$

For low blood soluble gases ($\lambda_{b:air} \leq 10$) [92] the measured concentration is dependent on the rates at which blood is pumped through the lungs and ventilation, specifically the ventilation-perfusion ratio, $\frac{V_A}{Q_c}$, where $C_{measured} = C_A$, meaning that low blood soluble VOCs must exchange completely in the alveoli.

Highly soluble VOCs ($\lambda_{b:air} > 10$) [92] tend to be less affected by changes in ventilation and perfusion, however, hydrophilic exhaled trace gases, such as acetone, interact with the water-like mucus membrane lining the conductive airways. The exhaled breath concentrations of these volatiles appear to dilute on their way up from the deeper respiratory track to the airway opening (dilution effect), consequently for these highly soluble volatiles the concentration measured in exhaled breath is different from the alveolar air concentration, $C_{measured} \neq C_A$. There is also evidence that, with highly soluble volatile compounds, gas exchange occurs in the airways rather than alveoli [94].

Some studies have been performed assessing the concentration profiles during exercise [66,113,115–117]. Recently, the influence of exercise on mouth-exhaled and nose-exhaled breath was further investigated [117]. Smith and co-workers reported significant increase of isoprene breath levels which are in agreement with previous findings [66].

An isoprene gas-exchange model (2) was developed and showed good fit to breath isoprene levels measured during exercise. Dependency of heartbeat rate and breath rate for isoprene breath

concentrations have been assessed, where isoprene levels were measured during exercise [113,116]. Isoprene concentrations showed drastic increase within the initials seconds of exercise [113,116], followed by a decline when heartbeat rate reached the maximum value and respiration rate increased, and lastly at the end of the exercise isoprene concentrations reached similar levels seen at the beginning. This means that the degree of blood-to-air partitioning of isoprene is very sensitive to heart rate. Such measurements demonstrates a relationship between breath rate volume (V*br*), heartbeat volume (HBV), Henry's law constant (*H*) and temperature (T), seen in Equation (2) [116].

$$C_{A0} = C_{V0} \cdot exp\left(-\frac{1}{HRT} \cdot \frac{Vbr}{HBV}\right) \quad (2)$$

For volatiles such as isoprene, with low solubility in blood and high volatility (Henry's law constant extremely low) a concentration gradient within the lungs is created and governed by the velocity of the bloodstream pumped through the lungs (proportional to heartbeat frequency) and the breathing rate. Namely, with increases in both heart rate and breathing rate, more efficient partitioning of isoprene to breath air is restored. This means that isoprene evaporates efficiently through the transport via the bloodstream to the lungs, hence, $C_{A0} \neq C_{V0}$, meaning that isoprene venous blood concentration entering the lungs, C_{V0}, is different from isoprene arterial blood concentration leaving the lungs, C_{A0}.

Moreover, measures taken during sleep showed enhanced blood isoprene concentration due to lower heartbeat rate achieved during the night [116].

5. Disease Diagnosis

5.1. Volatile Biomarkers

Current cancer detection methods include Computer tomography (CT) scanning, magnetic resonance imaging (MRI), as well as biopsies. However, several cancers such as lung cancer, colorectal cancer, bladder and prostate cancer are very difficult to detect at an early stage due to the lack of sensitivity of those methods. SIFT-MS and PTR-MS have, therefore, been used to determine whether they have potential for identification of possible cancer biomarkers. Other applications include diagnosis in liver disease, infectious diseases, food intolerances, and monitoring of diabetes.

5.1.1. Lung Cancer

Lung cancer is the most common cancer and it has one of the lowest survival outcomes of any cancer because over two-thirds of patients are diagnosed at a late stage when curative treatment is not possible. Methylated hydrocarbons are proposed for lung or breast cancer biomarkers [1].

Nowadays, it is well known that acetaldehyde is present in breath of healthy people [59] at a physiological mean level of about 22 ppb. Acetaldehyde above physiological levels in exhaled breath could have major clinical importance. However, these levels in breath may not be obvious in most of the cases. Acetaldehyde is an intermediate in the metabolism of ethanol in the liver, however, intake of alcohol will greatly elevate acetaldehyde levels in breath [63]. In addition, acetaldehyde can also be produced from cellular activity involving sugars.

Phillips and co-workers suggested, in a cross-sectional study, a combination of 22 VOCs potential markers of lung cancer, in breath samples of patients with and without lung cancer [4].

To support lung cancer studies, cells *in vitro* studies were performed in order to analyse the molecular emissions from cancer cells lines SK-MES and CALU-1 [118]. The experimental results showed that acetaldehyde is present in the headspace above cell cultures at levels significantly higher than physiological levels, a potential lung cancer biomarker. A contradicting publication reported by Filipiak and co-workers, using cell lines CALU-1 which volatiles were analysed by GC-MS, showed that compounds are not released but seem to be consumed by CALU-1 cells. These findings confirmed the existence of compounds that are either released or consumed by these cells [119]. Nevertheless, do the cell lines growing *in vitro* have similar characteristics to *in vivo* cancer cells? A 3D model has been proposed [120] in which the cells were cultured in 3D scaffolds composed of collagen type I hydrogels, compared to 2D models where cells are grown on surfaces such as plastic or glass. Quantification by SIFT-MS of cells lines headspace CALU-1 and non-malignant lung cells NL20 revealed that the amount of acetaldehyde released by both cell types grown in a 3D model is higher when compared to that of the same cells grown in 2D models.

5.1.2. Colorectal Cancer

Colorectal cancer has been attributed to individual genetic predisposition and environmental factors, including lifestyle and diet. Within lifestyle factors, elevated body mass index (BMI), obesity, and low physical activity are related to increased risk of colorectal cancer. It has been shown that diet can significantly influence and promote the growth of malignant colon cells, particularly, red meat intake, where protein is the major constituent leading to protein fermentation metabolites potentially carcinogenic and possible linked to colon cancer [121–124]. Therefore, diet may have complex effects on the generation of breath compounds [125]. The findings indicated that diets low in fat and rich in protein induced systemic ketosis leading to increased levels of acetone in breath [126]. Acetone may be reduced to isopropanol by hepatic alcohol dehydrogenase and, consequently, it may appear in the breath [125]. Carbohydrate fermentation by bacteria in the gut results in production of hydrogen, methane [127], carbon dioxide, and short chain fatty acids (SCFAs) mainly acetate, propionate and butyrate as the main non-gaseous fermentation end products. Early measurements of hydrogen in breath have been used to study carbohydrate absorption in the small intestine [128]. Similarly, measurements of methane in breath have been used to assess colonic bacterial metabolism [129]. Short chain fatty acids are assimilated by the host and used for the energy metabolism. Butyrate in particular [130] has been considered to have a protective role, protecting against colitis [130] and colorectal cancer [131]. Hence, bacteria have an important role on the generation of the majority of some compounds present in breath, such as hydrogen, hydrogen cyanide, aldehydes, and alkanes [125]. Recently, Schmidt and co-workers reported that breath hydrogen cyanide may rise following the consumption of food or drink [132]. Ammonia, isovaleric and isobutyric acid (BCFA), phenolics and hydrogen sulphide have been identified in breath as products of gastrointestinal bacterial fermentation.

Pentane, ethane and ethylene have also been identified in breath as products of increased lipid peroxidation [125].

Recently, Altomare and co-workers used GC–MS to analyse breath samples from patients with colorectal cancer, and concluded that the pattern of VOCs in patients suffering from colorectal cancer were different from that in healthy controls, particularly levels of some specific VOCs such as 1,3-dimethylbenzene, 1,2-pentadiene, cyclohexene and methylcyclohexene [133]. However, further studies are necessary to support this experiment.

Normal metabolism generates VOCs that may emanate from faecal matter. Human faecal flora comprises bacteria involved in colonic fermentation producing sulphur containing compounds, such as hydrogen sulphide, dimethyl disulphide, methyl disulphide, and dimethyl trisulphide, and these compounds are responsible for the specific odour of faecal matter [134]. Thus, analyses of human metabolites as end products of intestine may rely on faecal samples or on breath.

5.1.3. Breast Cancer

Breast cancer is accompanied by increased oxidative stress caused by lipid peroxidation of polyunsaturated fatty acids in membranes, producing alkanes and methylalkanes, such as 3-methylundecane, 6-methylpentadecane, and 2-methylpropane, among others, potential biomarkers as suggested by Phillips and co-workers. Hietanen and co-workers performed an initial case-control study where they analysed the breath of women with breast cancer and found increased concentrations of pentane [135]. Nevertheless, the majority of the investigations of volatile biomarkers potentially used in breast cancer diagnosis has been done by Phillips and co-workers. They performed a pilot study of breath VOCs in women with breast cancer. Breath samples were analysed by GC-MS and compared with abnormal mammograms and biopsies. The breath test distinguished between women with breast cancer and healthy volunteers with a sensitivity of 94.1% [136]. Their recent findings were consistent with previous studies, however the biochemical origin of volatile biomarkers of breast cancer remains speculative [8].

5.1.4. Liver Disease

Few studies have been performed in order to use breath analysis as a screening tool for liver disease diagnosis. Sulfur-containing compounds, such as dimethylsulfide, hydrogen sulphide, and mercaptans (e.g., methylmercaptan and ethylmercaptan) are proposed as liver cancer biomarkers [1].

Ammonia levels rise in the blood when the liver is unable to convert ammonia to urea. This may occur because of cirrhosis or severe hepatitis, however the majority of ammonia seen in mouth-exhaled breath is largely generated in the oral cavity [47]. Poor oral hygiene can be a confounding factor, because production of ammonia from urea may increase the ammonia levels in exhaled breath. Such a situation may be mitigated by mouth washing thoroughly with water before breath sampling [50]. Very little information is available about the possible use of VOCs in patients with liver cirrhosis. An initial study performed by Van den Velde and co-workers [137] showed discrimination between the cirrhotic group and healthy subjects. They analysed the breath of 50 patients with established liver cirrhosis. Recently, a pilot study using PTR-MS equipped with a time-of-flight mass analyzer was conducted in liver cirrhosis patients by sampling the breath of the subjects [138]. The authors were able to distinguish between healthy and disease subjects. They identified twelve different VOCs significantly different between cirrhotic and healthy subjects.

Hepatic encephalopathy [139] is a neuropsychiatric syndrome with symptoms varying depending on the severity of the condition. It results from the accumulation of compounds not cleared by the liver. Ammonia is known to be involved in hepatic encephalopathy [139], however attempts to use breath ammonia measurements for diagnosis have failed, probably because most exhaled ammonia is generated within the oral cavity by bacterial and/or enzymatic [47]. Additionally, some confounding factors have proven to be tricky in VOCs detection and quantification. For example, pulmonary gas exchange abnormalities can be present in patients with advanced liver disease, such as high cardiac output and abnormal dilation of pulmonary capillary vessels, leading to incorrect conclusions [139].

5.1.5. Infectious Diseases–Tuberculosis

Pulmonary tuberculosis (TB) is an infectious disease derived from the organism Mycobacterium tuberculosis. The primary detection technique is the Ziehl-Neelsen staining combined with microscopy. It only allows detection of pulmonary disease in an advanced stage, meaning that often the disease has already been transmitted to close contacts.

Breath analysis may offer a method for diagnosing pulmonary tuberculosis [140]. Some compounds potential biomarkers have been found by breath sampling, such as methyl phenylacetate, methyl *p*-anisate, methyl nicotinate, and *o*-phenylanisole [141]. However, contradictory studies presented different marker compounds, possible due to the fact that *Mycobacterium tuberculosis* is a slow growing organism. This means that if VOCs are produced they may be released or modified by the host, hence, they may well be present at low concentration and not be detected by the real-time analytical techniques. The best approach should be looking at clinical samples in order to get potential biomarkers of pulmonary tuberculosis [36].

5.1.6. Food Intolerances

Other breath analysis studies have been performed such as from people suffering from carbohydrate malabsorption, a condition in which the patients are unable to absorb or digest certain carbohydrates due to the lack of some intestinal enzymes, leading to bacterial sugar fermentation in the gut. Coeliac disease is an under-diagnosed autoimmune disease of the small intestine characterized by nutritional malabsorption, for which was conducted a preliminary investigation of the levels of alcohols in the breath of 10 patients with coeliac disease compared to that in 10 healthy controls using SIFT-MS [142]. No significant conclusions were drawn. Such conclusions are in agreement with recent findings performed by Aprea and co-workers [143], where real time breath analysis was performed in patients diagnosed with coeliac disease under gluten free diet. As expected, exhaled breath of patients with coeliac disease was similar to the exhaled breath of healthy people and no reliable marker was found.

A novel approach to the diagnosis of gastro-intestinal diseases was attempted by Lechner and co-workers using PTR-MS, through headspace screening of fluid obtained from the gut during colonoscopy and analysis of exhaled breath, either from healthy controls and patients suffering from inflammatory bowel disease (IBD) and irritable bowel syndrome (IBS) [144]. Fluid samples of patients with IBD showed enhanced peaks at m/z = 57 and m/z = 83, while no significant differences were found for IBS patients group. Further comparison of breath samples revealed increased concentration of ions at m/z = 31 and m/z = 77 in the IBS group. Lechner and co-workers suggested that the ions

detected at m/z = 31 and m/z = 77 probably represent protonated formaldehyde and protonated acetone with an attached water molecule respectively [144]. However, the use of m/z values for biomarker detection is unreliable, especially at higher m/z, as ions could represent any or many of a number of compounds. Recently, Dryahina and co-workers [145] reported pentane as potential biomarker of bowel disease, by analysing the breath of patients with Crohn's disease and ulcerative colitis and of healthy volunteers, in a pilot study using SIFT-MS.

6. Clinical Studies

6.1. Diabetes Mellitus

Centuries ago, John Gallo reported a compound in human breath that had the smell of decaying apples [146]. Nowadays, it is well known that the compound was principally acetone [147]. Acetone is produced by decarboxylation of acetoacetate and through dehydrogenation of isopropanol [61]. Diabetic patients exhibit increased concentrations in blood and urine of ketone bodies, acetone, acetoacetic acid, and beta-hydroxybutyric acid. Ketone bodies are produced by the liver during fatty acid metabolism, and are used as an energy source, instead of glucose, when glucose is not readily available [148].

Acetone is a highly soluble gas present as a major common breath metabolite in everyone. This volatile organic compound has been identified and quantified previously [40,57]. Elevated breath acetone levels were early associated with diabetes mellitus [61,149,150].

A few studies have been carried out, involving cohorts of patients with diabetes and healthy controls, in who parallel blood glucose levels have been determined. Biochemical changes that occur in the disease state (diabetic state) may be reflected in changes in the profile of VOCs in exhaled breath. The high inter-individual variability in breath acetone concentration is not well understood. Contributors may include diurnal variability, fasting status, diet, age, and gender.

Acetone levels were previously quantified for healthy volunteers, before and after ingesting 75 g of glucose in the fasting state [60]. The intake diet seems to significantly influence the levels in breath acetone. Measurements were taken following a ketogenic diet [126,151] in the exhaled breath of healthy individuals, and for a small group of individuals suffering from diabetes. Results have shown that breath acetone concentrations increased after ingestion of a ketogenic meal, or either following a low carbohydrate diet [151]. Smith and co-workers reported that breath acetone increases substantially during fasting when a change takes place from carbohydrate to fat metabolism [148]. Schwarz and co-workers, using PTR–MS, reported a variation on breath acetone with age and fasting state, but no statistically significant differences between gender or body-mass index (BMI) were found [152].

Glycaemic control is essential for management of diabetes. At the moment, blood analysis is the faster way to provide information/monitoring of glucose levels in patients with diabetes. However, studies amongst adults showed discomfort associated to blood sampling and needle phobia [153,154]. Only recently, breath analysis appears as a potential non-invasive method for monitoring glucose concentrations in blood [148,155,156]. Turner and co-workers have monitored the breath of eight patients with type 1 diabetes mellitus using a glucose clamp technique [156]. In all patients, the breath acetone declined linearly with blood glucose concentration. Hence, this study indicates that breath acetone does vary as glycaemia and/or metabolic status changes in type 1 diabetes.

Compounds present in exhaled breath displayed very strong correlations with glucose concentrations, in another study conducted in healthy and type 1 diabetic subjects [157]. Standard least squares regression was used on several subsets of exhaled gases to generate models to predict plasma glucose for each subject.

Little is currently known for type 2 diabetes, some measurements of breath acetone in type 2 diabetes were taken, nevertheless, with no significant results [158].

A recent study performed by Righettoni and co-workers reported the correlations between blood glucose of healthy volunteers and breath components, from portable gas sensors and PTR-MS equipped with a time-of-flight mass analyzer [159]. The relationship between the PTR-MS measurements of breath gases acetone, isoprene, ethanol and methanol, sensor response and the blood glucose level was studied. They reported a better correlation between blood glucose level and breath acetone for the overnight fasting (morning).

Breath composition during oral glucose tolerance tests was analysed by TD GC-MS in 16 subjects and correlated to blood glucose levels [160]. The glucose tolerance tests classified five of the subjects as diabetics, eight as affected by impaired glucose tolerance, and three as normoglycaemic. A clustering algorithm was used to differentiate individuals between groups based on blood glucose values at different times, and breath acetone concentrations. Acetone levels were generally higher in diabetics.

Dogs have been used to detect hypoglycaemic episodes in their diabetic owners through detecting breath or skin odour [161].

Isoprene has also been proposed as a potential indicator of diabetes [162]. However, several studies reported no apparent correlation between blood glucose and breath isoprene [58,163]. Methyl nitrate is also suggested to be correlated with blood glucose in insulin dependent diabetics, though the levels are reported to be lower than acetone or isoprene. Therefore, this might not be a useful compound for monitoring purposes [148].

Although the findings have pointed to acetone as potential biomarker of diabetes, there is no simple association of breath acetone concentration and diabetes. The issue lies in the fact that acetone generation is linked to lipolysis and blood glucose changes.

7. Biomarkers *vs*. Biomarker Profiles

The complex relationships between a number of different compounds and the presence or absence of a disease or condition indicates that perhaps volatile biomarkers profiling with bioinformatics is a more promising approach. A specific breath marker related to a specific disease is the ideal. However, this is unlikely to be the case for the majority of diseases or conditions, where it is more probable that a range of VOCs with varying concentrations will have to be used. By adopting a strategy of identifying patterns, rather than trying to identify individual VOCs, "breath fingerprinting" could provide a suitable and reliable method for discriminating between healthy and diseased states. This approach requires elaborate methods of data analysis, pattern recognition techniques, such as principal component analysis (PCA) and partial least squares discriminant analysis (PLSDA). Principal component analysis is a mathematical algorithm that reduces the dimensionality of the data. It accomplishes this reduction by identifying directions, called principal components, in which the variation in the data is maximal. Samples can be plotted, and visually assess similarities and differences

between samples, and determined whether samples can be grouped. Other multivariate methods also exist, such as PLSDA, or support vector machines (SVMs). All of these methods use whole profiles, and yet it is possible to identify individual components (e.g., compounds or ions), which are most responsible for the differences observed between groups (e.g., groups of samples positive or negative for a disease). Thus, "biomarkers" (which are not usually unique) may be identified in this way. Cross validation of the models is used to predict the classification capabilities on unknown objects. Hence, if there is a good correlation between the predicted and actual values this means that the model fits.

In clinical practice, biomarkers such as genes and proteins are identified and quantified in order to track the biochemistry within the body for a specific disease. However, clear quantification of VOCs is a harder job due to the difficulties in finding the biochemical pathways in the body for each metabolite. This will involve a close collaboration between clinicians and analytical chemists.

The need to understand the relationships between many variables makes multivariate analysis an inherently difficult subject. It is important to note that when the number of variables quickly overwhelms the number of samples, spurious correlations may be found [164]. Confounding variables have a real statistical correlation with the disease and a breath marker, leading to wrong conclusions. Confounding variables comprise environmental compounds, physiological parameters and even the sampling procedures [164]. The statistical technique used to control the influence of confounding variables is called Analysis of covariance (ANCOVA). Classification of the subjects into groups is achieved by discriminant analysis, cluster analysis, and propensity score analysis. Clustering attempts to find similarities among the subjects that were measured instead of among the measures that were made. For multiple dependent variables, in which two or more dependent variables are included, multivariate analysis of variance (MANOVA) and canonical correlation analysis are applied. Recently, Halbritter and co-workers [165] used MANOVA technique to discriminate according to whether the pregnant women had gestational diabetes mellitus, impaired glucose tolerance, or normal glucose tolerance, by means of analysing the women´s breath by PTR-MS and correlating it with the oral glucose tolerance test.

The success of proper statistical analysis is to have a good statistical validation, as well as trustworthy biological interpretation of the results.

8. Breath Test as a Clinical Diagnostic

Initial pilot exploratory studies aimed to establish the population distributions of the metabolites levels in breath of the healthy population [55]; to study the enhancement of breath metabolites by drug ingestion [166]; to investigate the influence of smoking on the breath metabolites [3,80]; among others. For such studies, the sample size of volunteers is not a restriction.

The concepts of standard error and confidence interval should be highly understood in order to determine the sample size of patients needed to develop a diagnostic test. The need of power calculations to calculate the minimum sample size required should be used. The sample size determines the amount of sampling error inherent in a test result. Effects are harder to detect in smaller samples leading to increased standard errors. Increasing sample size is often the easiest way to boost the statistical power of a test. A common recommendation by statisticians calls for ten times as many subjects as the number of independent variables. Great concern should be taken about sample size of

controls and subjects with the disease. However, several studies performed up to date used small control groups or/and small groups with the disease, due to the difficulty and cost of obtaining a significant sample size of subjects. Nevertheless, such strict rules are not required for essential pilot exploratory studies, as may be needed eventually. Moreover, the selection of appropriate controls is an important issue to take into account. An ideal control group should not have the disease in question, but be comparable to the diseased group, for instance have identical symptoms.

Until now, the clinical issues are far behind the analytical techniques, which have been proven to be extremely accurate and sensitive for trace gas analysis. As proven by the extensive studies of hydrogen cyanide (HCN) in relation to *Pseudomonas aeruginosa* infection, performed by Smith and co-workers by SIFT-MS [167]. They reported hydrogen cyanide as a volatile biomarker of *Pseudomonas aeruginosa* infection.

There is the need for identification of the relationships between biochemical pathways and disease; normal ranges and limiting concentrations for the breath VOCs for healthy subjects; and variations of VOCs concentrations with age, gender, and ethnicity. Along with identification of the minimum concentrations of biomarkers indicative of disease, the concentration of biomarkers, which identify the stage of disease, should be known as should the variations of abnormal concentrations with age, gender and ethnicity. Furthermore, clinical diagnostic tests using breath must be suitable for all categories of patients, namely children, adults, and elderly patients, and sensitivity and specificity be determined. Particular attention must be paid to patients with disabilities, ventilated patients, and asthmatic patients. The breath test must be as accurate as the existing screening tests, cheaper, quicker, and non-invasive to patients.

The clinical significance of the breath test will determine the applicability of it into clinical practice.

9. Concluding Remarks

Detection of volatile compounds in breath at trace concentrations can be an indicator of metabolic status, allowing identification of diseases in their early stages. Breath analysis has proved to be suited for the diagnosis of cancer, infectious diseases, food intolerances, and diabetes, among others. However, the complexity of VOCs in exhaled breath makes breath a difficult sample to analyze and the findings are not always consistent. Therefore, it has been difficult to reach an agreement regarding the identification of volatile biomarkers.

In order to measure the concentration of a volatile compound in exhaled breath, some understanding of the compound's exhalation physiology is necessary. Important factors, such as mouth generated volatiles, pulmonary gas exchange, contamination, *etc*., should not be neglected. To introduce breath analysis into clinical practice, sampling procedures have to be standardized and the metabolic pathways clearly understood. In addition, factors that are unrelated to disease, but capable of changing the concentration of a volatile compound in exhaled breath must be well understood in order to develop robust clinical tests. The findings lead to the suggestion that the creation of a unique diagnostic test would not be the best option. Instead, the creation of a diagnostic test according to the disease would be the best approach.

Acknowledgments

This work was supported by the EU (project 287382, PIMMS-Proton Ionization Molecular Mass Spectrometry) and funded by Marie Curie Actions—Initial Training Networks (ITN).

Author Contributions

Célia Lourenço planned and wrote the manuscript.
Claire Turner co-planned the manuscript and commented on and modified drafts.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Risby, T.H. Current status of clinical breath analysis. In *Breath Analysis for Clinical Diagnosis and Therapeutic Monitoring*; World Scientific: Tuck Link, Singapore, 2005; pp. 251–265.
2. Smith, D.; Španěl, P. Application of ion chemistry and the SIFT technique to the quantitative analysis of trace gases in air and on breath. *Int. Rev. Phys. Chem.* **1996**, *15*, 231–271.
3. Jordan, A.; Hansel, A.; Holzinger, R.; Lindinger, W. Acetonitrile and benzene in the breath of smokers and non-smokers investigated by proton transfer reaction mass spectrometry (PTR-MS). *Int. J. Mass Spectrom.* **1995**, *148*, L1–L3.
4. Phillips, M.; Gleeson, K.; Hughes, J.; Greenberg, J.; Cataneo, R.N.; Baker, L.; McVay, P. Volatile organic compounds in breath as markers of lung cancer: A cross-sectional study. *Lancet* **1999**, 353, 1930–1933.
5. Lacy Costello, B.; Amann, A.; Al-Kateb, H.; Flynn, C.; Filipiak, W.; Khalid, T.; Osborne, D.; Ratcliffe, N.M. A review of the volatiles from the healthy human body. *J. Breath Res.* **2014**, *8*, 1–29.
6. Smith, D.; Španěl, P. The challenge of breath analysis for clinical diagnosis and therapeutic monitoring. *Analyst* **2007**, *132*, 390–396.
7. Phillips, M.; Cataneo, R.N.; Ditkoff, B.A.; Fisher, P.; Greenberg, J.; Gunawardena, R.; Kwon, C.S.; Tietje, O.; Wong, C. Prediction of breast cancer using volatile biomarkers in the breath. *Breast Cancer Res. Treat.* **2006**, *99*, 1–3.
8. Phillips, M.; Cataneo1, R.N.; Saunders, C.; Hope, P.; Schmitt, P.; Wai, J. Volatile biomarkers in the breath of women with breast cancer. *J. Breath Res.* **2010**, *4*, 1–8.
9. Smith, D.; Španěl, P.; Herbig, J.; Beauchamp, J. Mass spectrometry for real-time quantitative breath analysis. *J. Breath Res.* **2014**, doi:10.1088/1752-7155/8/2/027101.
10. Borsdorf, H.; Eiceman, G.A. Ion mobility spectrometry: Principles and applications. *Appl. Spec. Rev.* **2006**, *41*, 323–375.
11. Ruzsanyi, V.; Mochalski, P.; Schmid, A.; Wiesenhofer, H.; Klieber, M.; Hinterhuber, H.; Amann, A. Ion mobility spectrometry for detection of skin volatiles. *J. Chromatogr. B* **2012**, *911*, 84–92.

12. Westhoff, M.; Litterst, P.; Freitag, L.; Urfer, W.; Bader, S.; Baumbach, J.I. Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: Results of a pilot study. *Thorax* **2009**, *64*, 744–748.
13. Perl, T.; Carstens, E.; Hirn, A.; Quintel, M.; Vautz, W.; Nolte, J.; Jünger, M. Determination of serum propofol concentrations by breath analysis using ion mobility spectrometry. *Br. J. Anaesth.* **2009**, *103*, 822–827.
14. McCurdy, M.; Bakhirkin, Y.; Wysocki, G.; Lewicki, R.; Tittel, F.K. Recent advances of laser-spectroscopy- based techniques for applications in breath analysis. *J. Breath Res.* **2007**, *1*, 1–12.
15. Vaittinen, O.; Manfred Schmidt, F.; Metsala, M.; Halonen, L. Exhaled breath biomonitoring using laser spectroscopy. *Curr. Anal. Chem.* **2013**, *9*, 463–475.
16. Adams, N.G.; Smith, D. The Selected Ion Flow Tube (SIFT): A technique for studying thermal energy ion-neutral reactions. *Int. J. Mass. Spectrom. Ion Phys.* **1976**, *21*, 349–359.
17. Smith, D.; Adams, N.G. The Selected Ion Flow Tube (SIFT): Studies of ion-neutral reactions. *Adv. Atom. Mol. Phys.* **1987**, *24*, 1–49.
18. Smith, D. The ion chemistry of interstellar clouds. *Chem. Rev.* **1992**, *92*, 1473–1485.
19. Španěl, P.; Smith, D. Selected ion flow tube: A technique for quantitative trace gas analysis of air and breath. *Med. Biol. Eng.Comput.* **1996**, *34*, 409–419.
20. Smith, D.; Španěl, P. The novel selected-ion flow tube approach to trace gas analysis of air and breath. *Rapid. Commun. Mass Spectrom.* **1996**, *10*, 1183–1198.
21. Španěl, P.; Rolfe, P.; Rajant, B.; Smith, D. The selected ion flow tube (SIFT)—A novel technique for biological monitoring. *Ann. Occup. Hyg.* **1996**, *40*, 615–626.
22. Španěl, P.; Smith, D. Selected ion flow tube mass spectrometry for on-line trace gas analysis in biology and medicine. *Eur. J. Mass. Spectrom.* **2007**, *13*, 77–82.
23. Smith, D.; Španěl, P. Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass. Spectrom. Rev.* **2005**, *24*, 661–700.
24. Smith, D.; Sovová, K.; Španěl, P. A selected ion flow tube study of the reactions of $H3O^+$, NO^+ and $O2^{+\bullet}$ with seven isomers of hexanol in support of SIFT-MS. *Int. J. Mass Spectrom.* **2012**, 25–30.
25. Španěl, P.; Smith, D. Progress in SIFT-MS: Breath Analysis and other applications. *Mass. Spectrom. Rev.* **2011**, 30, 236–267.
26. Španěl, P.; Smith, D. Advances in on-line absolute trace gas analysis by SIFT-MS. *Curr. Anal. Chem.* **2013**, *9*, 525–539.
27. Lindinger, W.; Hansel, A.; Jordan, A. On-line monitoring of volatile organic compounds at pptv levels by means of Proton-Transfer-Reaction Mass Spectrometry (PTR-MS) Medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Processes* **1998**, *173*, 191–241.
28. Jordan, A.; Haidacher, S.; Hanel, G.; Hartungen, E.; Herbig, J.; Märk, L.; Schottkowsky, R.; Seehauser, H.; Sulzer, P.; Märk, T.D. An online ultra-high sensitivity Proton-transfer-reaction mass-spectrometer combined with switchable reagent ion capability (PTR + SRI-MS). *Int. J. Mass Spectrom.* **2009**, *286*, 32–38.

29. Herbig, J.; Müller, M.; Schallhart, S.; Titzmann, T.; Graus, M.; Hansel, A. On-line breath analysis with PTR-TOF. *J. Breath Res.* **2009**, *3*, 1–10.
30. Jordan, A.; Haidacher, S.; Hanel, G.; Hartungen, E.; Märk, L.; Seehauser, H.; Schottkowsky, R.; Sulzer, P.; Märk, T.D. A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS). *Int. J. Mass Spectrom.* **2009**, *286*, 122–128.
31. Beauchamp, J.; Herbig, J.; Dunkl, J.; Singer, W.; Hansel, A. On the performance of proton-transfer-reaction mass spectrometry for breath-relevant gas matrices. *Meas. Sci. Technol.* **2013**, *24*, 1–13.
32. Hewitt, C.N.; Hayward, S.; Tani, A. The application of proton transfer reaction-mass spectrometry (PTR-MS) to the monitoring and analysis of volatile organic compounds in the atmosphere. *J. Environ. Monit.* **2003**, *5*, 1–7.
33. Herbig, J.; Amann, A. Proton transfer reaction-mass spectrometry applications in medical research. *J. Breath Res.* **2009**, *3*, 1–2.
34. Shurmer, H.; Fard, A.; Barker, J.; Bartlett, P.; Dodd, G.; Hayat, U. Development of an electronic nose. *Phys. Technol.* **1987**, *18*, 170–176.
35. Li, B.; Sauvé, G.; Iovu, M.; Jeffries-EL, M.; Zhang, R.; Cooper, J.; Santhanam, S.; Schultz, L.; Revelli, J.; Kusne, A.; *et al.* Volatile organic compound detection using nanostructured copolymers. *Nano Lett.* **2006**, *6*, 1598–1602.
36. Amann, A.; Smith, D. *Volatile Biomarkers Non-Invasive Diagnosis in Physiology and Medicine*; Elsevier: Oxford, UK, 2013.
37. Righettoni, M.; Tricoli, A.; Gass, S.; Schmid, A.; Amann, A.; Pratsinis, S. Breath acetone monitoring by portable Si:WO_3 gas sensors. *Anal. Chim. Acta* **2012**, *738*, 69–75.
38. Pauling, L.; Robinson, A.B.; Teranishi, R.; Cary, P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc. Nat. Acad. Sci. USA* **1971**, *68*, 2374–2376.
39. Miekisch, W.; Schubert, J.K.; Noeldge-Schomburg, G. Diagnostic potential of breath analysis—focus on volatile organic compounds. *Clin. Chim. Acta* **2004**, *347*, 25–39.
40. Smith, D.; Turner, C.; Španěl, P. Volatile metabolites in the exhaled breath of healthy volunteers: Their levels and distributions. *J. Breath Res.* **2007**, *1*, 1–12.
41. Buszewski, B.; Kesy, M.; Ligor, T.; Amann, A. Human exhaled air analytics: Biomarkers of diseases. *Biomed. Chromatogr.* **2007**, *21*, 553–566.
42. Phillips, M.; Greenberg, J.; Sabas, M. Alveolar gradient of pentane in normal human breath. *Free Radic. Res.* **1994**, *20*, 333–337.
43. Schubert, J.K.; Miekisch, W.; Birken, T.; Geiger, K.; Noeldge-Schomburg, G. Impact of inspired substance concentrations on the results of breath analysis in mechanically ventilated patients. *Biomarkers* **2005**, *10*, 138–152.
44. Španěl, P.; Dryahina, K.; Smith, D. A quantitative study of the influence of inhaled compounds on their concentrations in exhaled breath. *J. Breath Res.* **2013**, *7*, 1–10.
45. Phillips, M.; Cataneo, R.N.; Greenberg, J.; Munawar, M.I.; Nachnani, S.; Samtani, S. Pilot study of a breath test for volatile organic compounds associated with oral malodor: Evidence for the role of oxidative stress. *Oral Dis.* **2005**, *11*, 32–34.
46. Wang, T.; Pysanenko, A.; Dryahina, K.; Španěl, P. Analysis of breath, exhaled via the mouth and nose, and the air in the oral cavity. *J. Breath Res.* **2008**, *2*, 1–13.

47. Smith, D.; Wang, T.; Pysanenko, A.; Španěl, P. A selected ion flow tube mass spectrometry study of ammonia in mouth- and nose-exhaled breath and in the oral cavity. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 783–789.
48. Pysanenko, A.; Spaněl, P.; Smith, D. A study of sulfur-containing compounds in mouth- and nose-exhaled breath and in the oral cavity using selected ion flow tube mass spectrometry. *J. Breath Res.* **2008**, *2*, 1–13.
49. Khalid, T.Y.; Saad, S.; Greenman, J.; Costello, B.L.; Probert, C.S.J.; Ratcliffe, N.M. Volatiles from oral anaerobes confounding breath biomarker discovery. *J. Breath Res.* **2013**, *7*, 1–12.
50. Španěl, P.; Turner, C.; Wang, T.; Bloor, R.; Smith, D. Generation of volatile compounds on mouth exposure to urea and sucrose: Implications for exhaled breath analysis. *Physiol. Meas.* **2006**, *27*, N7–N17.
51. Amann, A.; Miekisch, W.; Pleil, J.; Risby, T.; Schubert, J.K. Methodological issues of sample collection and analysis of exhaled breath. *Exhal. Biomark*; European Respiratory Society: Plymouth, UK, 2010; pp. 96–107.
52. Ulanowska, A.; Kowalkowski, T.; Hrynkiewicz, K.; Jackowskic, M.; Buszewskia, B. Determination of volatile organic compounds in human breath for Helicobacter pylori detection by SPME-GC/MS. *Biomed. Chromatogr.* **2011**, *25*, 391–397.
53. Smith, D.; Pysanenko, A.; Španěl, P. Kinetics of ethanol decay in mouth- and nose-exhaled breath measured on-line by selected ion flow tube mass spectrometry following varying doses of alcohol. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1066–1074.
54. Španěl, P.; Smith, D. On-line measurement of the absolute humidity of air, breath and liquid headspace samples by selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 563–569.
55. Diskin, A.M.; Španěl, P.; Smith, D. Time variation of ammonia, acetone, isoprene and ethanol in breath: A quantitative SIFT-MS study over 30 days. *Physiol. Meas.* **2003**, *24*, 107–119.
56. Turner, C.; Španěl, P.; Smith, D. A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol. Meas.* **2006**, *27*, 637–648.
57. Turner, C.; Španěl, P.; Smith, D. A longitudinal study of ammonia, acetone and propanol in the exhaled breath of 30 subjects using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol. Meas.* **2006**, *27*, 321–337.
58. Turner, C.; Španěl, P.; Smith, D. A longitudinal study of breath isoprene in healthy volunteers using selected ion flow tube mass spectrometry (SIFT-MS). *Physiol. Meas.* **2006**, *27*, 13–22.
59. Turner, C.; Španěl, P.; Smith, D. A longitudinal study of ethanol and acetaldehyde in the exhaled breath of healthy volunteers using selected-ion flow-tube mass spectrometry. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 61–68.
60. Turner, C.; Parekh, B.; Walton, C.; Španěl, P.; Smith, D.; Evans, M. An exploratory comparative study of volatile compounds in exhaled breath and emitted by skin using selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 526–532.
61. Kalapos, M.P. On the mammalian acetone metabolism: From chemistry to clinical implications. *Biochim. Biophys. Acta* **2003**, *1621*, 122–139.

62. Lindinger, W.; Taucher, J.; Jordan, A.; Hansel, A.; Vogel, W. Endogenous production of methanol after the consumption of fruit. *Alcohol Clin. Exp. Res.* **1997**, *21*, 939–943.
63. Smith, D.; Wang, T.; Španěl, P. On-line, simultaneous quantification of ethanol, some metabolites and water vapour in breath following the ingestion of alcohol. *Physiol. Meas.* **2002**, *23*, 477–489.
64. Hyspler, R.; Crhova, S.; Gasparic, J.; Zadak, Z.; Cizkova, M.; Balasova, V. Determination of isoprene in human expired breath using solid-phase microextraction and gas chromatography-mass spectrometry. *J. Chromatogr. B* **2000**, *739*, 183–190.
65. Lirk, P.; Bodrogi, P.; Raifer, H.; Greiner, K.; Ulmer, H.; Rieder, J. Elective haemodialysis increases exhaled isoprene. *Nephrol. Dial. Transpl.* **2003**, *18*, 937–941.
66. King, J.; Koc, H.; Unterkofler, K.; Mochalski, P.; Kupferthaler, A.; Teschl, G.; Teschl, S.; Hinterhuber, H.; Amann, A. Physiological modeling of isoprene dynamics in exhaled breath. *J. Theor. Biol.* **2010**, *267*, 626–637.
67. Enderby, B.; Lenney, W.; Brady, M.; Emmett, C.; Španěl, P.; Smith, D. Concentrations of some metabolites in the breath of healthy children aged 7–18 years measured using selected ion flow tube mass spectrometry (SIFT-MS). *J. Breath Res.* **2009**, *3*, 1–11.
68. Pöschl, G.; Seitz, H.K. Alcohol and cancer. *Alcohol Alcsm* **2004**, *39*, 155–165.
69. Taucher, J.; Hansel, A.; Jordan, A.; Fall, R.; Futrell, J.H.; Lindinger, W. Detection of isoprene in expired air from human subjects using proton-transfer-reaction mass spectrometry. *Rapid. Commun. Mass Spectrom.* **1997**, *11*, 1230–1234.
70. Lechner, M.; Moser, B.; Niederseer, D.; Karlseder, A.; Holzknecht, B.; Fuchs, M.; Colvin, S.; Tilg, H.; Rieder, J. Gender and age specific differences in exhaled isoprene levels. *J. Respir. Physiol. Neurobiol.* **2006**, *154*, 478–483.
71. Španěl, P.; Dryahina, K.; Smith, D. Acetone, ammonia and hydrogen cyanide in exhaled breath of several volunteers aged 4–83 years. *J. Breath Res.* **2007**, *1*, 1–4.
72. Smith, D.; Španěl, P.; Enderby, B.; Lenney, W.; Turner, C.; Davies, S.J. Isoprene levels in the exhaled breath of 200 healthy pupils within the age range 7–18 years studied using SIFT-MS. *J. Breath Res.* **2010**, *4*, 1–7.
73. Španěl, P.; Smith, D. Selected ion flow tube–mass spectrometry: detection and real-time monitoring of flavours released by food products. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 585–596.
74. Winkler, K.; Herbig, J.; Kohl, I. Real-time metabolic monitoring with proton transfer reaction mass spectrometry. *J. Breath Res.* **2013**, *7*, 1–8.
75. Taucher, J.; Hansel, A.; Jordan, A.; Lindinger, W. Analysis of compounds in human breath after ingestion of garlic using proton-transfer-reaction mass spectrometry. *J. Agric. Food Chem.* **1996**, *44*, 3778–3782.
76. Diskin, A.M.; Španěl, P.; Smith, D. Increase of acetone and ammonia in urine headspace and breath during ovulation quantified using selected ion flow tube mass spectrometry. *Physiol. Meas.* **2003**, *24*, 191–199.
77. Smith, D.; Ismail, K.; Diskin, A.M.; Chapman, G.; Magnay, J.; Španěl, P.; O'Brien, S. Increase of acetone emitted by urine in relation to ovulation. *Acta Obstet. Gynecol.* **2006**, *85*, 1008–1011.

78. Abbott, S.M.; Elder, J.B.; Španěl, P.; Smith, D. Quantification of acetonitrile in exhaled breath and urinary headspace using selected ion flow tube mass spectrometry. *Int. J. Mass Spectrom.* **2003**, *228*, 655–665.
79. Buszewski, B.; Ulanowska, A.; Ligor, T.; Denderz, N.; Amann, A. Analysis of exhaled breath from smokers, passive smokers and non-smokers by solid-phase microextraction gas chromatography/mass spectrometry. *Biomed. Chromatogr.* **2009**, *23*, 551–556.
80. Prazeller, P.; Karl, T.; Jordan, A.; Holzinger, R.; Hansel, A.; Lindinger, W. Quantification of passive smoking using proton-transfer-reaction mass spectrometry. *Int. J. Mass Spectrom.* **1998**, *178*, L1–L4.
81. Habib, M.P.; Clements, N.C.; Garewal, H.S. Cigarette smoking and ethane exhalation in humans. *Am. J. Respir. Crit. Care Med.* **1995**, *151*, 1368–1372.
82. Kushch, I.; Schwarz, K.; Schwentner, L.; Baumann, B.; Dzien, A.; Schmid, A.; Unterkofler, K.; Gastl, G.; Španěl, P.; Smith, D.; *et al.* Compounds enhanced in a mass spectrometric profile of smokers' exhaled breath *vs.* non-smokers as determined in a pilot study using PTR-MS. *J. Breath Res.* **2008**, *2*, 1–26.
83. Sandberg, A.; Skold, C.M.; Grunewald, J. ; Eklund, A.; Wheelock, A.M. Assessing recent smoking status by measuring exhaled carbon monoxide levels. *PLoS One* **2011**, *6*, 1–7.
84. Beauchamp, J. Inhaled today, not gone tomorrow: Pharmacokinetics and environmental exposure of volatiles in exhaled breath. *J. Breath Res.* **2011**, *5*, 1–14.
85. Steerenberg, P.A.; Nierkens, S.; van Loveren, H.; van Amsterdam, J.G.C. A simple method to sample exhaled NO not contaminated by ambient NO from children and adults in epidemiological studies. *Nitric Oxide* **2000**, *4*, 168–174.
86. Pleil, J.; Lindstrom, A.B. Sample timing and mathematical considerations for modeling breath elimination of volatile organic compounds. *Risk Anal.* **1998**, *18*, 585–602.
87. Society, A.T. Recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide in adults and children-1999. *Am. J. Respir. Crit. Care Med.* **1999**, *160*, 2104–2117.
88. Society, A.T. ATS/ERS Recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am. J. Respir. Crit. Care Med.* **2005**, *171*, 912–930.
89. Horváth, I.; Hunt, J.; Barnes, P.J. Exhaled breath condensate: Methodological recommendations and unresolved questions. *Eur. Respir. J.* **2005**, *26*, 523–548.
90. Beauchamp, J.; Pleil, J. Simply breath-taking? Developing a strategy for consistent breath sampling. *J. Breath Res.* **2013**, *7*, 1–3.
91. Gilchrist, F.J.; Razavi, C.; Webb, A.K.; Jones, A.M.; Spaněl, P.; Smith, D.; Lenney, W. An investigation of suitable bag materials for the collection and storage of breath samples containing hydrogen cyanide. *J. Breath Res.* **2012**, *6*, 1–7.
92. Anderson, J.; Babb, A.; Hlastala, M.P. Modeling soluble gas exchange in the airways and Alveoli. *Ann. Biomed. Eng.* **2003**, *31*, 1402–1422.
93. Levitt, M.; Ellis, C.; Furne, J. Influence of method of alveolar air collection on results of breath tests. *Dig. Dis. Sci.* **1998**, *43*, 1938–1945.

94. Anderson, J.; Hlastala, M. Breath tests and airway gas exchange. *Pulm. Pharmacol. Ther.* **2007**, *20*, 112–117.
95. Herbig, J.; Titzmann, T.; Beauchamp, J.; Kohl, I.; Hansel, A. Buffered end-tidal (BET) sampling—A novel method for real-time breath-gas analysis. *J. Breath Res.* **2008**, *2*, 1–9.
96. Di Francesco, F.; Loccioni, C.; Fioravanti, M.; Russo, A.; Pioggia, G.; Ferro, M.; Roehrer, I.; Tabucchi, S.; Onor, M. Implementation of Fowler's method for end-tidal air sampling. *J. Breath Res.* **2008**, *2*, 1–9.
97. Fowler, W.S. Lung function studies: II. The respiratory dead space. *Am. J. Physiol.* **1948**, *154*, 405–416.
98. Dubowski, K.M. Biological aspects of breath-alcohol analysis. *Clin. Chem.* **1974**, *20*, 294–299.
99. Schubert, J.K.; Spittler, K.H.; Braun, G.; Geiger, K.; Guttmann, J. CO_2-controlled sampling of alveolar gas in mechanically ventilated patients. *J. Appl. Physiol.* **2001**, *90*, 486–492.
100. Birken, T.; Schubert, J.; Miekisch, W.; Nöldge-Schomburg, G. A novel visually CO_2 controlled alveolar breath sampling technique. *Technol. Health Care* **2006**, *14*, 499–506.
101. Filipiak, W.; Filipiak, A.; Ager, C.; Wiesenhofer, H.; Amann, A. Optimization of sampling parameters for collection and preconcentration of alveolar air by needle traps. *J. Breath Res.* **2012**, *6*, 1–19.
102. Cope, K.A.; Watson, M.T.; Foster, W.M.; Sehnert, S.S.; Risby, T.H. Effects of ventilation on the collection of exhaled breath in humans. *J. Appl. Physiol.* **2004**, *96*, 1371–1379.
103. Anderson, J.; Lamm, W.; Hlastala, M. Measuring airway exchange of endogenous acetone using a single-exhalation breathing maneuver. *J. Appl. Physiol.* **2006**, *100*, 880–889.
104. O'Hara, M.E.; O'Hehir, S.; Green, S.; Mayhew, C.A. Development of a protocol to measure volatile organic compounds in human breath: A comparison of rebreathing and on-line single exhalations using proton transfer reaction mass spectrometry. *Physiol. Meas.* **2008**, *29*, 309–330.
105. King, J.; Unterkofler, K.; Teschl, S.; Mochalski, P.; Koc, H.; Hinterhuber, H.; Amann, A. A modeling-based evaluation of isothermal rebreathing for breath gas analyses of highly soluble volatile organic compounds. *J. Breath Res.* **2012**, *6*, 1–10.
106. Mochalski, P.; King, J.; Unterkofler, K.; Amann, A. Stability of selected volatile breath constituents in Tedlar, Kynar and Flexfilm sampling bags. *Analyst* **2013**, *138*, 1405–1418.
107. Beauchamp, J.; Herbig, J.; Gutmann, R. On use of Tedlar bags for breath-gas sampling and analysis. *J. Breath Res.* **2008**, *2*, 1–18.
108. Mochalski, P.; Wzorek, B.; Sliwka, I.; Amann, A. Suitability of different polymer bags for storage of volatile sulphur compounds relevant to breath analysis. *J. Chromatogr. B* **2009**, *877*, 189–196.
109. Van Harreveld, A. Odor concentration decay and stability in gas sampling bags. *J. Air Waste Manage. Assoc.* **2003**, *53*, 51–60.
110. Groves, W.; Zellers, E. Investigation of organic vapor losses to condensed water vapor in Tedlar bags used for exhaled-breath sampling. *Am. Ind. Hyg. Assoc. J.* **1996**, *57*, 257–263.
111. Trefz, P.; Rösner, L.; Hein, D.; Schubert, J.K.; Miekisch, W. Evaluation of needle trap micro-extraction and automatic alveolar sampling for point-of-care breath analysis. *Anal. Bioanal. Chem.* **2013**, *405*, 3105–3115.

112. Greger, R.; Windhorst, U. Pulmonary Gas Exchange. In *Comprehensive Human Physiology*; Springer-Verlag: Berlin, Germany, 1996; Volume 2., pp. 2037–2049.
113. Schubert, R.; Schwoebel, H.; Mau-Moeller, A.; Behrens, M.; Fuchs, P.; Sklorz, M.; Schubert, J.K.; Bruhn, S.; Miekisch, W. Metabolic monitoring and assessment of anaerobic threshold by means of breath biomarkers. *Metabolomics* **2012**, *8*, 1069–1080.
114. King, J.; Unterkofler, K.; Teschl, G.; Teschl, S.; Koc, H.; Hinterhuber, H.; Amann, A. A mathematical model for breath gas analysis of volatile organic compounds with special emphasis on acetone. *J. Math. Biol.* **2011**, *63*, 959–999.
115. King, J.; Unterkofler, K.; Teschl, S.; Mochalski, P.; Koc, H.; Hinterhuber, H.; Amann, A. Isoprene and acetone concentration profiles during exercise on an ergometer. *J. Breath Res.* **2009**, *6*, 1–11.
116. Karl, T.; Prazeller, P.; Mayr, D.; Jordan, A.; Rieder, J.; Fall, R.; Lindinger, W. Human breath isoprene and its relation to blood cholesterol levels: New measurements and modeling. *J. Appl. Physiol.* **2001**, *91*, 762–770.
117. Smith, D.; Chippendale, W.E.T.; Dryahina, K.; Španěl, P. SIFT-MS analysis of nose-exhaled breath; mouth contamination and the influence of exercise. *Curr. Anal. Chem.* **2013**, *9*, 565–575.
118. Smith, D.; Wang, T.; Sulé-Suso, J.; Španěl, P.; Haj, A.E. Quantification of acetaldehyde released by lung cancer cells *in vitro* using selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 845–850.
119. Filipiak, W.; Sponring, A.; Mikoviny, T.; Ager, C.; Schubert, J.K.; Miekisch, W.; Amann, A.; Troppmair, J. Release of volatile organic compounds (VOCs) from the lung cancer cell line CALU-1 *in vitro*. *Cancer Cell Int.* **2008**, *8*, 1–11.
120. Rutter, A.; Chippendale, W.E.T.; Yang, Y.; Španěl, P.; Smith, D.; Sulé-Suso, J. Quantification by SIFT-MS of acetaldehyde released by lung cells in a 3D model. *Analyst* **2013**, *138*, 91–95.
121. Cummings, J.H.; Bingham, S.A. Dietary fibre, fermentation and large bowel cancer. *Cancer Surv.* **1987**, *6*, 601–621.
122. Van Munster, I.P.; Nagengast, F.M. The role of carbohydrate fermentation in colon-cancer prevention. *Scand. J. Gastroenterol. Suppl.* **1993**, *28*, 80–86.
123. Windey, K.; Preter, V.; Verbeke, K. Relevance of protein fermentation to gut health. *Mol. Nutr. Food Res.* **2012**, *56*, 184–196.
124. Larsson, S.; Wolk, A. Meat consumption and risk of colorectal cancer: A meta-analysis of prospective studies. *Int. J. Cancer* **2006**, *119*, 2657–2664.
125. Ajibola, O.A.; Smith, D.; Španěl, P.; Ferns, G.A.A. Effects of dietary nutrients on volatile breath metabolites. *J. Nutr. Sci.* **2013**, *2*, 1–15.
126. Musa-Veloso, K.; Likhodii, S.; Cunnane, S. Breath acetone is a reliable indicator of ketosis in adults consuming ketogenic meals. *Am. J. Clin. Nutr.* **2002**, *76*, 65–70.
127. Mckay, L.; Eastwood, M.A.; Brydon, M.G. Methane excretion in man-A study of breath, flatus, and faeces. *Gut* **1985**, *26*, 69–74.
128. Bond, J.H.; Levitt, M.D. Use of pulmonary hydrogen (H_2) measurements to quantitate carbohydrate malabsorption: Study of partially gastrectomized patients. *J. Clin. Invest.* **1972**, *51*, 1219–1225.

129. Lacy Costello, B.; Ledochowski, M.; Ratcliffe, N.M. The importance of methane breath testing: A review. *J. Breath Res.* **2013**, *7*, 1–8.
130. Pryde, S.E.; Duncan, S.H.; Hold, G.L.; Stewart, C.S.; Flint, H.J. The microbiology of butyrate formation in the human colon. *FEMS Microbiol. Lett.* **2002**, *217*, 133–139.
131. McIntyre, A.; Gibson, P.R.; Young, G.P. Butyrate production from dietary fibre and protection against large bowel cancer in a rat model. *Gut* **1993**, 34, 386–391.
132. Schmidt, F.M.; Metsälä, M.; Vaittinen, O.; Halonen, L. Background levels and diurnal variations of hydrogen cyanide in breath and emitted from skin. *J. Breath Res.* **2011**, *5*, 1–10.
133. Altomare, D.F.; Lena, M.D.; Porcelli, F.; Trizio, L.; Travaglio, E.; Tutino, M.; Dragonieri, S.; Memeo, V.; Gennaro, G. Exhaled volatile organic compounds identify patients with colorectal cancer. *Br. J. Surg.* **2013**, *100*, 144–150.
134. Moore, J.G.; Jessop, L.D.; Osborne, D.N. Gas-chromatographic and mass-spectrometric analysis of the odor of human feces. *Gastroenterology* **1987**, *93*, 1321–1329.
135. Hietanen, E.; Bartsch, H.; Bériziat, J.C.; Camus, A.M.; McClinton, S.; Eremin, O.; Davidson, L.; Boyle, P. Diet and oxidative stress in breast, colon and prostate cancer patients: A case-control study. *Eur. J. Clin. Nutr.* **1994**, *48*, 575–586.
136. Phillips, M.; Cataneo, R.N.; Ditkoff, B.; Fisher, P.; Greenberg, J.; Gunawardena, R.; Kwon, C.S.; Rahbari-Oskoui, F.; Wong, C. Volatile markers of breast cancer in the breath. *Breast J.* **2003**, *9*, 184–191.
137. Van den Velde, S.; Nevens, S.; van Hee, P.; van Steenberghe, D.; Quirynen, M. GC-MS analysis of breath odor compounds in liver patients. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2008**, *875*, 344–348.
138. Morisco, F.; Aprea, E.; Lembo, V.; Fogliano, V.; Vitaglione, P.; Mazzone, G.; Cappellin, L.; Gasperi, F.; Masone, S.; Domenico de Palma, G.; *et al.* Rapid "breath-print" of liver cirrhosis by proton transfer reaction time-of-flight mass spectrometry. A pilot study. *PLoS One* **2013**, *8*, 1–9.
139. Khalid, T.Y.; Lacy Costello, B.; Ewen, R.; White, P.; Stevens, S.; Gordon, F.; Collins, P.; McCune, A.; Shenoy, A.; Shetty, S.; *et al.* Breath volatile analysis from patients diagnosed with harmful drinking, cirrhosis and hepatic encephalopathy: A pilot study. *Metabolomics* **2013**, *9*, 938–948.
140. Phillips, M.; Cataneo, R.N. ; Condose, R.; Ring Ericksond, G.; Greenberga, J.; Bombardie, V.; Munawar, M.; Tietje, O. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis* **2007**, *87*, 44–52.
141. Syhre, M.; Manning, L.; Phuanukoonnon, S.; Harino, P.; Chambers, S.T. The scent of Mycobacterium tuberculosis–Part II breath. *Tuberculosis* **2009**, *89*, 263–266.
142. Hryniuk, A.; Ross, B.M. A preliminary investigation of exhaled breath from patients with celiac disease using selected ion flow tube mass spectrometry. *J. Gastrointestin. Liver Dis.* **2010**, *19*, 15–20.
143. Aprea, E.; Cappellin, L.; Gasperi, F.; Morisco, F.; Lembo, V.; Rispo, A.; Tortora, R.; Vitaglione, P.; Caporaso, N.; Biasioli, F. Application of PTR-TOF-MS to investigate metabolites in exhaled breath of patients affected by coeliac disease under gluten free diet. *J. Chromatogr. B* **2014**, doi:10.1016/j.jchromb.2014.02.015.

144. Lechner, M.; Colvin, H.P.; Ginzel, C.; Lirk, P.; Rieder, J.; Tilg, H. Headspace screening of fluid obtained from the gut during colonoscopy and breath analysis by proton transfer reaction-mass spectrometry: A novel approach in the diagnosis of gastro-intestinal diseases. *Int. J. Mass Spectrom.* **2005**, *243*, 151–154.
145. Dryahina, K.; Španěl, P.; Pospíšilová, V.; Sovová, K.; Hrdlička, L.; Machková, M.; Smith, D. Quantification of pentane in exhaled breath, a potential biomarker of bowel disease, using selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 1983–1992.
146. Crofford, O.B.; Mallard, R.E.; Winton, R.E.; Rogers, N.L.; Jackson, J.C.; Keller, U. Acetone in breath and blood. *Trans. Am. Clin. Climatol. Assoc.* **1977**, *88*, 128–139.
147. Reichard, G.A., Jr.; Skutches, C.L.; Hoeldtke, R.D.; Owen, O.E. Acetone Metabolism in Humans during diabetic ketoacidosis. *Diabetes* **1986**, *35*, 668–674.
148. Smith, D.; Španěl, P.; Fryer, A.; Hanna, F.; Ferns, G. Can volatile compounds in exhaled breath be used to monitor control in diabetes mellitus? *J. Breath Res.* **2011**, *5*, 1–8.
149. Manolis, A. The diagnostic potential of breath analysis. *Clin. Chem.* **1983**, *29*, 5–15.
150. Deng, C.; Zhang, J.; Yu, X.; Zhang, W.; Zhang, X. Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization. *J. Chromatogr. B* **2004**, *810*, 269–275.
151. Španěl, P.; Dryahina, K.; Rejskova, A.; Chippendale, W.E.T.; Smith, D. Breath acetone concentration; bilogical variability and the influence of diet. *Physiol. Meas.* **2011**, *32*, 23–31.
152. Schwarz, K.; Pizzini, A.; Arendack, B.; Zerlauth, K.; Filipiak, W.; Schmid, A.; Dzien, A.; Neuner, S.; Lechleitner, M.; Scholl-Bürgi, S.; *et al*. Breath acetone—aspects of normal physiology related to age and gender as determined in a PTR-MS study. *J. Breath Res.* **2009**, *3*, 1–9.
153. Simmons, J.; Mcfann, K.; Brown, A.; Rewers, A.; Follansbee, D.; Temple-Trujillo, R.; Klingensmith, G. Reliability of the diabetes fear of injecting and self-testing questionnaire in pediatric patients with type 1 diabetes. *Diabetes Care* **2007**, *30*, 987–988.
154. Hamilton, J.G. Needle phobia-A neglected diagnosis. *J. Fam. Pract.* **1995**, *41*, 169–175.
155. Guo, D.; Zhang, D.; Li, N.; Zhang, L.; Yang, J. Diabetes identification and classification by means of a breath analysis system. In Proceedings of the Medical Biometrics Second International Conference, ICMB 2010; Springer: Berlin, Germany, 2010; pp. 52–63.
156. Turner, C.; Walton, C.; Hoashi, S.; Evans, M. Breath acetone concentration decreases with blood glucose concentration in type I diabetes mellitus patients during hypoglycaemic clamps. *J. Breath Res.* **2009**, *3*, 1–6.
157. Minh, T.; Oliver, S.; Ngo, J.; Flores, R.; Midyett, R. ; Meinardi, S.; Carlson, M.K.; Rowland, F.S.; Blake, D.R.; Galassetti, P.R. Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects. *Am. J. Physiol. Endocrinol. Metab.* **2011**, *300*, 1166–1175.
158. Storer, M.; Dummer, M.; Lunt, H.; Scotter, J.; McCartin, F.; Cook, J.; Swanney, M.; Kendall, D.; Logan, F.; Epton, M. Measurement of breath acetone concentrations by selected ion flow tube mass spectrometry in type 2 Diabetes. *J. Breath Res.* **2011**, *5*, 1–5.
159. Righettoni, M.; Schmid, A.; Amann, A.; Pratsinis, S.E. Correlations between blood glucose and breath components from portable gas sensors and PTR-TOF-MS. *J. Breath Res.* **2013**, *7*, 1–9.

160. Ghimenti, S.; Tabucchi, S.; Lomonaco, T.; di Francesco, F.; Fuoco, R.; Onor, M.; Lenzi, S.; Trivella, M.G. Monitoring breath during oral glucose tolerance tests. *J. Breath Res.* **2013**, *7*, 1–7.
161. Wells, D.; Lawson, S.; Siriwardena, A.N. Canine responses to hypoglycemia in patients with type 1 diabetes. *J. Altern. Complement. Med.* **2008**, *14*, 1235–1241.
162. Salerno-Kennedy, R.; Cashman, K.D. Potential applications of breath isoprene as a biomarker in modern medicine: A concise overview. *Wien. Klin. Wochenschr.* **2005**, *117*, 180–186.
163. O'Hara, M.E.; Clutton-Brock, T.H.; Green, S.; Mayhew, C.A. Endogenous volatile organic compounds in breath and blood of healthy volunteers: Examining breath analysis as a surrogate for blood measurements. *J. Breath Res.* **2009**, *3*, 1–10.
164. Miekisch, W.; Herbig, J.; Schubert, J.K. Data interpretation in breath biomarker research: Pitfalls and directions. *J. Breath Res.* **2012**, *6*, 1–10.
165. Halbritter, S.; Fedrigo, M.; Hollriegl, V.; Szymczak, W.; Maier, J.M.; Ziegler, A.G.; Hummel, M. Human breath gas analysis in the screening of gestational diabetes mellitus. *Diabetes Technol. Therapeut.* **2012**, *14*, 917–925
166. Harrison, G.R.; Critchley, A.D.; Mayhew, C.A.; Thompson, J.M. Real-time breath monitoring of propofol and its volatile metabolites during surgery using a novel mass spectrometric technique: A feasibility study. *Br. J. Anaesth.* **2003**, *91*, 797–799.
167. Smith, D.; Španěl, P.; Gilchrist, F.J.; Lenney, W. Hydrogen cyanide, a volatile biomarker of Pseudomonas aeruginosa infection. *J. Breath Res.* **2013**, *7*, 1–13.

Appendix C

PLOS | ONE

RESEARCH ARTICLE

Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer

Claire A Batty[1☯], Michael Cauchi[2☯], Célia Lourenço[1☯], John O Hunter[3☯], Claire Turner[1☯*]

1 Dept. Life, Health & Chemical Sciences, The Open University, Walton Hall, Milton Keynes, United Kingdom, **2** Centre for Biomedical Engineering, Cranfield University, Cranfield, Bedfordshire, United Kingdom, **3** Gastroenterology Research Dept., Addenbrooke's Hospital, Cambridge, United Kingdom

☯ These authors contributed equally to this work.
* claire.turner@open.ac.uk

OPEN ACCESS

Citation: Batty CA, Cauchi M, Lourenço C, Hunter JO, Turner C (2015) Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer. PLoS ONE 10(6): e0130301. doi:10.1371/journal.pone.0130301

Academic Editor: Mathias Chamaillard, INSERM, FRANCE

Received: January 14, 2015

Accepted: May 19, 2015

Published: June 18, 2015

Data Availability Statement: All relevant data are within the paper and Supporting Information files.

Funding: Engineering and Physical Sciences Research Council, EPSRC, Grant EP/L011395/1 (http://www.epsrc.ac.uk/). Funding for CT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared no competing interests exist.

Abstract

Diagnosis of colorectal cancer is an invasive and expensive colonoscopy, which is usually carried out after a positive screening test. Unfortunately, existing screening tests lack specificity and sensitivity, hence many unnecessary colonoscopies are performed. Here we report on a potential new screening test for colorectal cancer based on the analysis of volatile organic compounds (VOCs) in the headspace of faecal samples. Faecal samples were obtained from subjects who had a positive faecal occult blood sample (FOBT). Subjects subsequently had colonoscopies performed to classify them into low risk (non-cancer) and high risk (colorectal cancer) groups. Volatile organic compounds were analysed by selected ion flow tube mass spectrometry (SIFT-MS) and then data were analysed using both univariate and multivariate statistical methods. Ions most likely from hydrogen sulphide, dimethyl sulphide and dimethyl disulphide are statistically significantly higher in samples from high risk rather than low risk subjects. Results using multivariate methods show that the test gives a correct classification of 75% with 78% specificity and 72% sensitivity on FOBT positive samples, offering a potentially effective alternative to FOBT.

Introduction

Colorectal cancer is the fourth most common cancer in men and the 3rd most common cancer in women. There is significant international variation in the distribution of this cancer [1] and it is often regarded as a disease of Western industrialised countries. This strongly suggests that environmental factors may play a major role in the aetiology [2], and the highest incidence rates of colorectal cancer are located in North America, Europe and Oceania [1,3].

Despite the higher incidence of this cancer, mortality has been found to have decreased in 13 out of 29 countries [1]. The decrease is generally considered to be due to improvements in treatment, screening and earlier detection and was seen in longstanding economically stable

countries such as the USA, Australia, France and the UK. Even better screening and detection may improve this further.

Screening Programmes and Strategies

Many countries now choose to screen for this disease as a matter of course to obtain earlier diagnosis. There is growing evidence that screening asymptomatic people who are at an average risk of colorectal cancer means cancers are generally detected earlier, are more 'curable' and result in an overall reduction in mortality [4]. It is thought that 70–90% of cases of colorectal cancer arise from premalignant (adenomatous) polyps. They often have a stalk which consists of healthy tissue and allows them to be removed simply and completely by endoscopic snaring which enhances the need for adequate screening [5].

In the UK, much effort has been placed on setting up an adequate screening programme. On the advice of the National Screening Committee, the UK Department of Health carried out a demonstration pilot to test the feasibility of a national screening programme for colorectal cancer [6]; these were then expanded to further pilots and proved successful in their aim of bringing colorectal cancer screening into the UK's National Health Service (NHS) [7]. It is now a nationally rolled out protocol where the NHS offer free screening to all men and women aged 60–69 in the UK every two years, and on request can be sent to anyone over 70 years of age [8].

In both the USA and the UK, the means of screening subjects are generally similar with initial screening being stool/faecal based tests, and then secondary structural based tests. In both the US and UK, the chosen faecal based tests is the faecal occult blood test (FOBT) and then the more invasive structural test is the colonoscopy/sigmoidoscopy [6, 7, 9, 10, 11].

The FOBT is designed to detect occult (i.e. not accompanied by discernible symptoms or signs) blood in the stool and works on the basis of the fact that large polyps and the actual colorectal cancers in the colon and rectum tend to bleed. The test requires the sampler to smear a card with faeces twice after a bowel movement, then repeat this for a further two bowel movements, making a total of six windows covered on the card. This is then sent in a hygienically sealed envelope back to the relevant clinic and is tested. A positive result is deemed to be when any of the windows return as 'positive' [12,13].

The benefits of this test include its relatively low invasiveness, low initial cost and the fact it requires few specialised resources. It is also possible to do this test at home easily, making it perfect for a large scale screening programme. It has also been proven to help reduce the incidence of colorectal cancer [4, 11, 14, 15]. However, the test is not specific for human haemoglobin and also does not take into account blood that may be from other sources, i.e. haemorrhoids [4] and peptic ulcers. Another problem is that the sensitivity and specificity of the test is variable. Its sensitivity has been reported at around 40% for cancer and only 24% for the detection of advanced adenoma or *in situ* carcinoma [16]. If the FOBT is used as a screening procedure for the general population, the predictive value of a positive test is no more than 5% to 10%. This means most patients who test positive will require an expensive, uncomfortable, confirmatory diagnostic procedure such as colonoscopy when in fact they do not have colorectal cancer [17]. Although FOBT is the most commonly used screening test, it does not offer the optimum requisites of a screening test, due to its lack of sensitivity and specificity. Another screening test is for abnormal DNA in a faecal sample. This abnormal DNA arises from mutations occurring in cancer or adenoma cells, which then may be shed into the faeces. Imperiale et al [18] compared this method with the FOBT and found that in a limited study, it proved to be more sensitive than FOBT with similar specificity.

Use of biomarkers in screening

The human gut has a microbiome with a large bacterial population which can be commensal, but has also been shown to cause some detrimental effects. Mounting evidence has shown a relationship between infective agents and some human cancers. Certain mucosal associated bacterial species play an important role in the pathogenesis of colorectal cancer [19]. Metabolite profiling of volatile organic compounds (VOCs) of human colon cell lines including normal, and human cell carcinoma, offer not only biochemical phenotyping of normal and neoplastic colon tissue, but also differences in the volatile metabolome in different disease stages in comparison to the volatile metabolome of non-malignant colon epithelial cell lines [20,21].

The 'scent' of disease is becoming more prominent; and canine scent detection, i.e. use of dogs, has shown that the dogs are able to differentiate between samples from cancer patients and those without cancer. This suggests and reinforces that there are different 'volatile compounds' between cancer and non-cancer [22,23].

It was also reported that the pattern of breath VOCs differed between healthy controls and those diagnosed with colorectal cancer. Analysis showed an initial specificity of 83% with a sensitivity of 86% and an overall accuracy of 85%. When this then went through further validation testing on 25 subjects, 19 were correctly identified leading to 76% accuracy. This is encouraging for a more accurate screening test [24].

The characteristic patterns of VOCs in faeces have been reported for the many different causes of diarrhoea. As a potentially fast and convenient method, it opens up a new area for use as a non-invasive diagnostic tool. It can be performed repeatedly, can be applied to children including neonates as samples are easy to collect, and to patients with severe disease where more invasive procedures are not easily possible [25,26].

In this research, the difference in volatile profiles of high and low risk colorectal samples is investigated to see whether VOC analysis has potential for screening or diagnosis of colorectal cancer.

Materials and Methods

Sample acquisition

The cancer samples in this study were collected through the National NHS Bowel Cancer Screening Scheme. This service is offered nationally every 2 years to all males and females between the age of 60 and 69. Subjects received an invitation letter explaining the programme along with an information leaflet from the programme hub centre. Approximately 1 week later, the FOBT is sent out to the subject with step-by-step instructions for completing the test at home and then how to send the sample back to the hub lab. In the test, participants are sent out a cardboard envelope containing three windows on which may be smeared a thin sample of faeces. In the kit there is a spatula for this purpose together with clear instructions. The test is called Haemoccult by Beckman Coulter. A sample is taken on three separate days and once complete, the participant then seals the cardboard envelope and sends it back to the laboratory. The test card is then processed at the hub lab and the results were returned to the subject within 2 weeks.

Subjects who tested positive with FOBT screening and then accepted a colonoscopy were invited to participate in this research. These patients are requested to produce a further stool sample to bring with them in a wide plastic sterile container provided by the hospital. It was these samples which we studied.

After the colonoscopy, the patients were stratified into risk groups based on their histopathology. In total there were 7 classes as illustrated in Table 1.

In this study, 31 samples were obtained from the 'normal' or low risk group (class 1) and 31 samples were obtained from the high risk groups (class 5 and 6). However, all samples were from individuals who had screened positive using the FOB test. After samples were obtained from patients, they were stored at -80°C until processing by selected ion flow tube mass spectrometry (SIFT-MS) as described below. The resulting data were then processed using multivariate statistical techniques.

Ethics statement

Samples were taken under the National NHS bowel cancer screening scheme. Favourable ethical opinion was obtained for subsequent inclusion of suitable subjects in this research by Cambridgeshire 2 REC, REF 08/H0308/13, and written informed consent obtained from each subject and consent form then kept securely. All samples were anonymised prior to analysis and only the cancer status was known to the authors.

Sample Preparation

Human faecal samples were removed from the -80°C freezer and 5g sub samples were weighed out. Each sub sample was placed into a Nalophan bag and sealed with a Swagelokconnector and tube. The bag was filled with hydrocarbon free air and placed in the incubator for 45 minutes.

SIFT-MS analysis

Full details of how SIFT-MS may be used to analyse trace gases and volatile organic compounds may be found elsewhere [27], however, a brief explanation is warranted here. In SIFT-MS, precursor ions (H_3O^+, NO^+ and O_2^+) are generated in a microwave discharge and the chosen ion is selected by a quadrupole mass filter. The selected ion is then injected into a fast flowing helium carrier gas, and down a flow tube. A sample is then introduced into the flow tube, and the precursor ion reacts with the trace gases and volatile organic compounds in the sample. The precursor and product ions in the carrier gas are separated in a second quadrupole mass spectrometer and subsequently counted in a detector. Data may be obtained through scanning a spectrum at a user-defined range of mass-to-charge ratio (m/z) values and quantification is carried out using a kinetics database stored in the instrument.

After incubation, sample bags in the incubator were connected to the SIFT-MS via the heated sampling line, and then analysed using each of the three precursor ions available in SIFT-MS and in the same order of H_3O^+ then NO^+ and finally O_2^+. The m/z range was set from 10 to 140, and the total time for analysis using each precursor was 30 seconds.

Table 1. Risk Class Applied to Histological Results.

Class	Detail	Number of samples
1	Normal, nothing of concern found = **low risk**	31
2	Hyperplastic polyp	0
3	Adenoma—Polyp less than 10mm diameter	0
4	Adenoma—polyp more than 10mm diameter	0
5	High Grade Adenoma = **high risk**	31
6	Adenocarcinoma = **high risk**	
99	Other (for example IBD)	0

doi:10.1371/journal.pone.0130301.t001

Blank samples of background air were also analysed. Data were then transferred to the computer to be analysed for the range of compounds present in each sample.

Statistical analysis

Data (in file S1 Table) were analysed using appropriate univariate and multivariate statistical techniques. Multivariate data analysis was performed by custom-built scripts written in MATLAB R2011a (MathWorks Inc., Nattick, USA) using functions from the PLS Toolbox (version 3.5, EigenVector Research Inc., USA).

The univariate technique involved the non-parametric Mann-Whitney U test analysis which was used because the data were not normally distributed. The analysis was performed using Excel and was chosen as this test does not assume normal distribution or that the variances of two populations are equal. It is also a useful technique for smaller sample groups and is based on the comparison of each observation from the first group with each observation from the second group. The data from each group are then individually compared together. It is also a useful technique for semi-quantitative data and has good probability of producing statistically significant results [28].

Prior to multivariate data analysis, data generated by the H_3O^+, NO^+ and O_2^+ precursor ions were combined into one large dataset. The intensity values at the range of m/z values within the data were normalised against the intensity values of the H_3O^+ precursor ions (m/z value of 19). The m/z values pertaining to the known adducts (isotopologues) of the precursor ions were removed; these had the following m/z values: 19, 21, 30, 32, 34, 37, 39, 48, 55, 57, 66, 73, 75, and 91. Finally, m/z values pertaining to data columns consisting only of zeros were also removed.

Exploratory data analysis using principal components analysis [29] revealed no outlying samples.

Following feature selection via either the parametric student *t* test [30] or the non-parametric Wilcoxon T test [31], multivariate classification via partial least squares discriminant analysis (PLS-DA) was performed [32,33]. This is a supervised pattern recognition technique in which the computer is trained to recognise patterns in the data that will help distinguish between low and high risk patients. To ensure a robust and confident result was attained, a two-step bootstrapping process was employed in which the first step deduces an optimal model, and the second step evaluates the model. This was repeated 150 times to attain an average performance of all models created.,. The statistical significance of the classification accuracy was determined by performing permutation testing [33] which involved randomising the class assignation 300 times, and for each random assignation, the two-step bootstrapping process was performed. The statistical z-test [34] was employed to determine the significance ($p < 0.05$).

Results and Discussion

Univariate Analysis

Initial comparisons using the Mann-Whitney U test analysis and the H_3O^+ precursor ion showed two ions to be statistically significant. These were m/z 35 (Fig 1) and m/z 90 (Fig 2)

When using the NO^+ precursor ion, no ions were found to be statistically significant. Finally the O_2^+ data were analysed using Mann-Whitney techniques and two ions were found to be statistically significant. These were m/z 62 and m/z 94. Fig 3 illustrates the significant m/z values found and their corresponding identities and values.

One of the tentatively identified VOCs is hydrogen sulphide. As a low molecular weight compound, it is easier to identify with SIFT-MS due to fewer potential other overlapping ions,

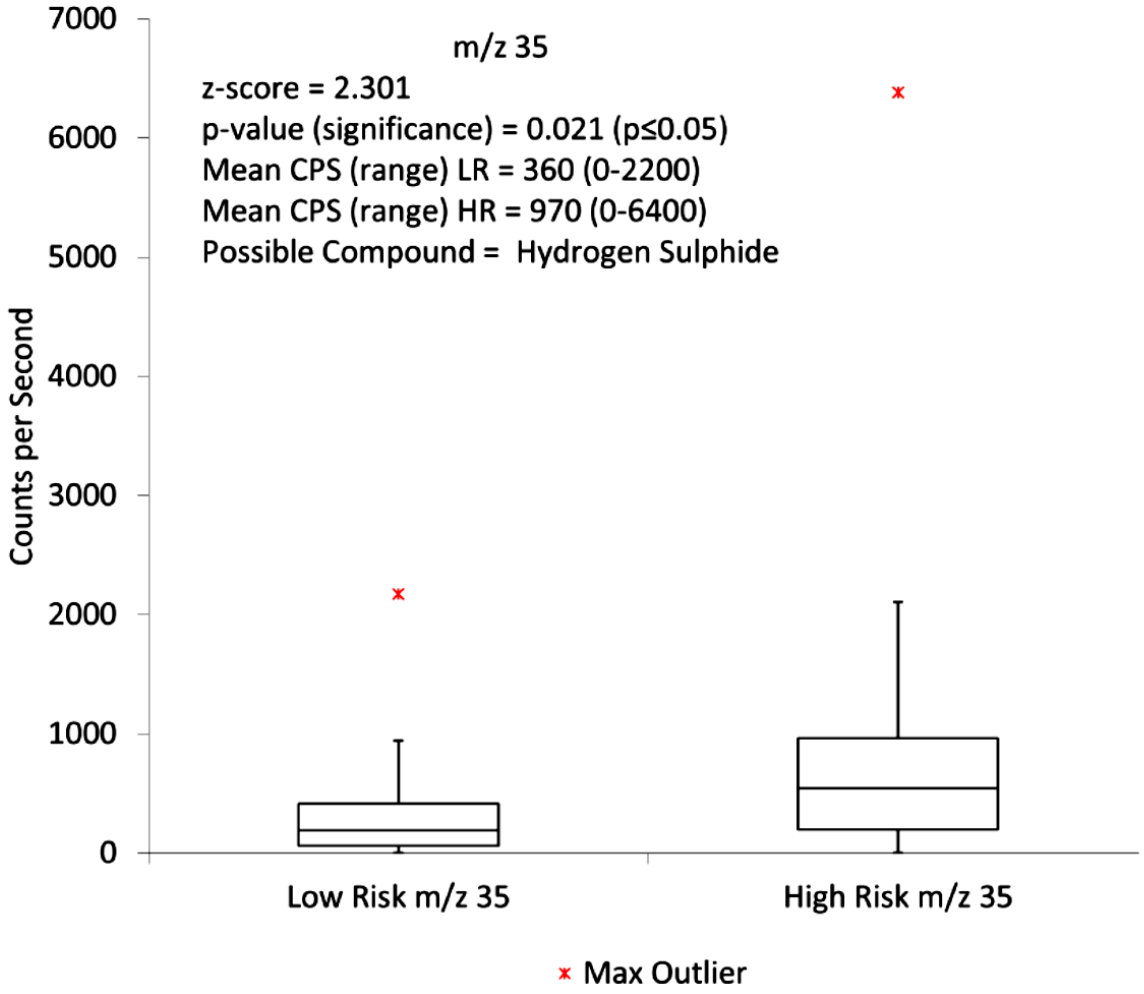


Fig 1. Box and whisker plot showing m/z 35 when comparing high risk to low risk faecal headspace samples for colorectal cancer with the H_3O^+ precursor ion. CPS = counts per second, LR = low risk, HR = high risk.

doi:10.1371/journal.pone.0130301.g001

and there is a reasonable likelihood that m/z35 is indeed H_2S. Hydrogen sulphide has been widely discussed in regards to the gut and colorectal cancer. It is now well known that hydrogen sulphide is endogenous and is produced by both enzymatic reaction and microbiota in the gut [35]. The concentration in human faeces is seen to be at 'normal' levels when it is within the range of 1.0–2.4 mmol/kg [36]. At lower concentrations hydrogen sulphide produces a positive biological effect but at higher levels, it becomes toxic. Interestingly it does not require a specific receptor for intracellular signalling and binds to the haem molecule. The important effect of hydrogen sulphide is its binding to cytochrome c oxidase [35, 37, 38] which is the crucial mechanism responsible for its toxicity. Binding of oxygen to cytochrome c oxidase is

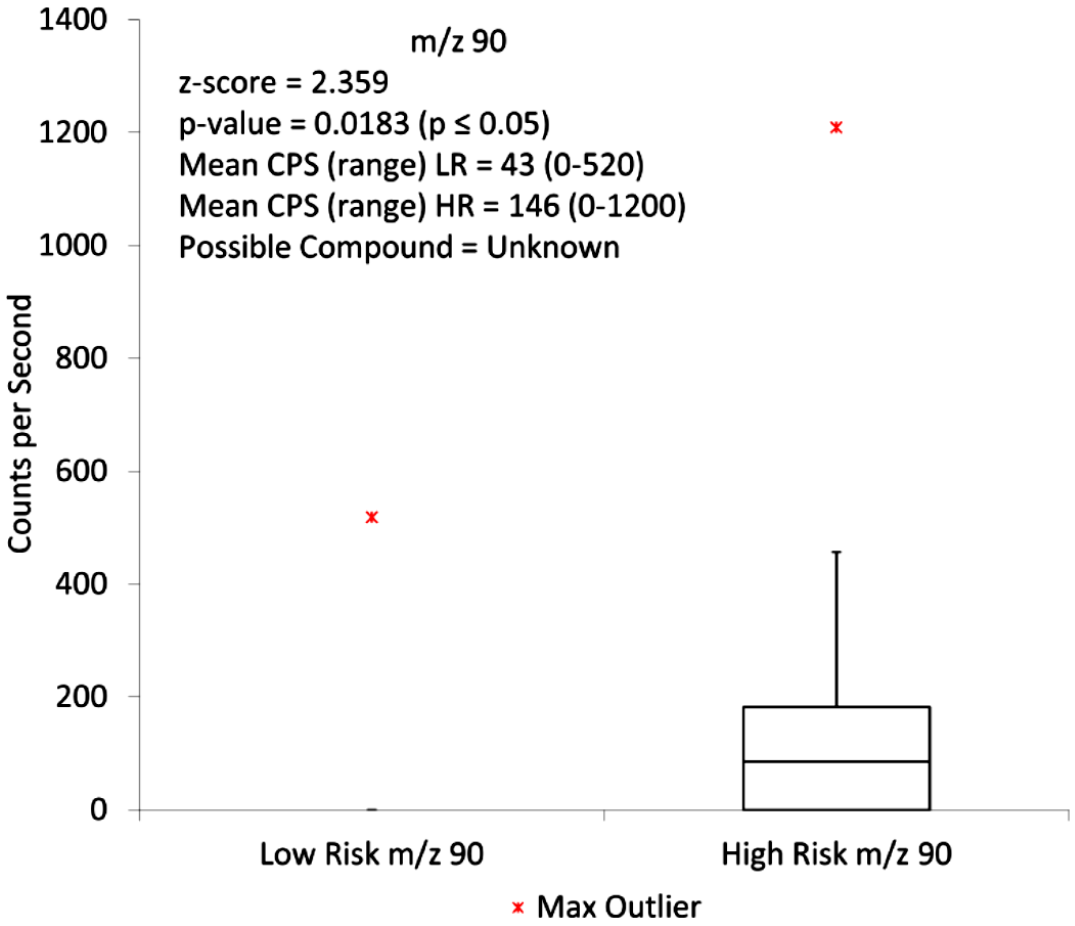


Fig 2. Box and whisker plot showing the distribution of low risk and high risk samples for m/z 90 using the H_3O^+ precursor ion. CPS = counts per second, LR = low risk, HR = high risk.

doi:10.1371/journal.pone.0130301.g002

inhibited by hydrogen sulphide non-competitively. This inhibition causes a reduction in cellular adenosine triphosphate (ATP) which subsequently has a direct effect on the ATP sensitive potassium channels. The activation of these channels plays a major role in the balance of the biological effects of hydrogen sulphide. Persistent colonisation by sulphate reducing bacteria and presence of hydrogen sulphide in the gut and faeces have been shown to be present in patients with ulcerative colitis as well as colorectal cancer [39, 40, 41] and have been implicated in the role of generating DNA damage at the genomic level [34]. Failure of colonocytes to differentiate appropriately may mean that they are more exposed to hydrogen sulphide in the lumen. The resulting effect of this may be inflammation in the gut and cell death as seen in ulcerative colitis. Continual irritation in this manner may well then lead to the genetic changes that ultimately cause colorectal cancer [42]. The mechanism by which this damage is thought to occur is via the disruption of the balance of apoptosis, proliferation and differentiation in

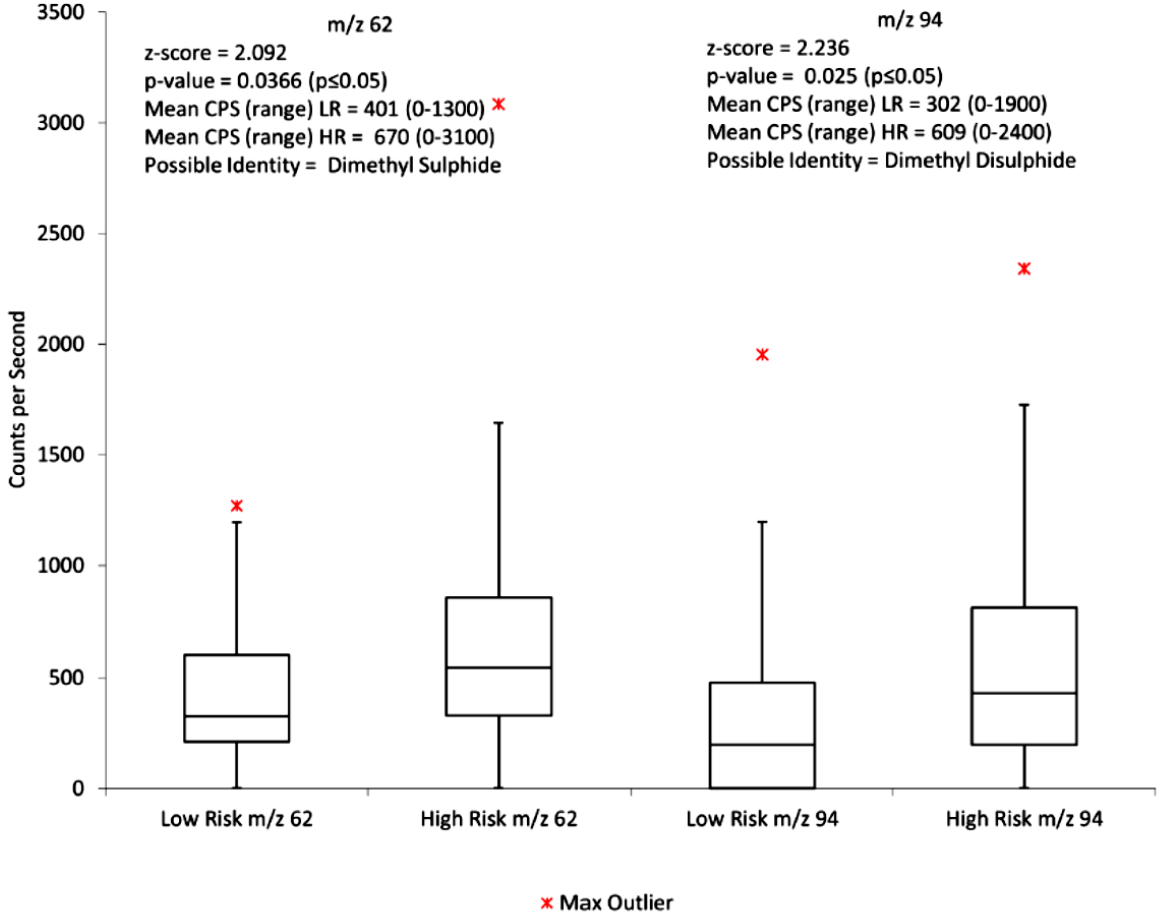


Fig 3. Box and whisker plot showing the distribution of low risk and high risk samples for m/z 62 and m/z 94 using the O_2^+ precursor ion.
CPS = counts per second, LR = low risk, HR = high risk.

doi:10.1371/journal.pone.0130301.g003

the intestinal epithelium. These data indicate that hydrogen sulphide could be an important factor in the progress, development and diagnosis of colorectal cancer.

Multivariate Analysis

On analysis with PLS-DA, the best resulting data were achieved by combining data sets using H_3O^+, NO^+ and O_2^+ using PLS-DA following feature selection via the student *t* test (STT). Here an overall classification accuracy of 75% was achieved which had a specificity of 78% and a sensitivity of 72%.

PLS-DA following feature selection with the Wilcoxon T test (WTT) also produced a promising result with a classification accuracy of 75%, a specificity of 79%, and a sensitivity of 70%. The respective distributions are shown in Fig 4.

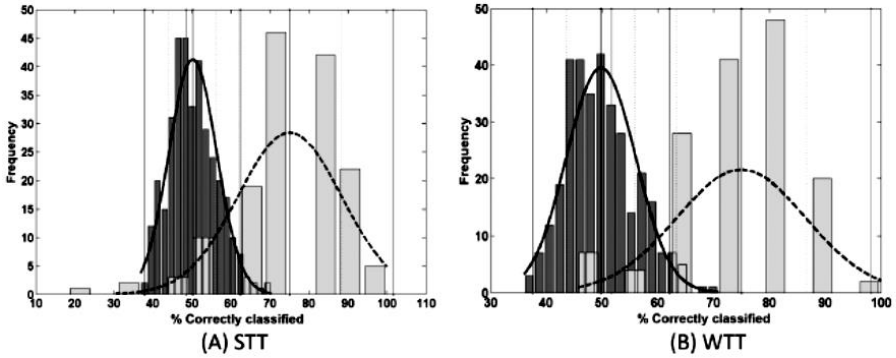


Fig 4. Permutation test distribution (dark grey (n = 300)) and the distribution of the bootstrapping classification accuracies (light grey (n = 150)) via PLS-DA following feature selection via STT (A) and WTT (B) for low versus high risk for all precursor ion datasets combined.

doi:10.1371/journal.pone.0130301.g004

The two-tailed z-tests produced calculated p-values of $< 1.0 \times 10^{-6}$ for STT (Fig 4A) and WTT (Fig 4B) at $\alpha = 0.05$. As the calculated p-values are much less than α, the Null Hypothesis is rejected which therefore proves that there is a significant difference between the two groups, i.e. low and high risk.

Further analysis of the optimal PLS-DA model resulted in the generation of PLS-DA scores and loadings. The scores demonstrate how the high and low risk groups have been distributed and are shown in Fig 5 for STT and WTT. The loadings permit the individual m/z values to be identified that contribute to the distinction between the low and high risk samples. These values are shown in Table 2.

Fig 5 shows that the distinction between the low and high risk is captured in the Latent variable 2 (LV2) axis for both STT and WTT which is supported by the fact that the optimum number of latent variables (LVs) which produced the optimum models in both instances was 2.

Interpreting the results

Following the use of multivariate data analysis on the SIFT-MS data, the analytical method and the technique applied to it showed a 75% correct classification to put samples into either their high risk or low risk group, when all precursor ions were combined in the analysis. This along with a specificity of 78% and a sensitivity of 72% means that overall this method is better at screening for colorectal cancer than the FOBT. The data proved to be particularly of interest, due to the fact the colorectal samples were taken after an initial positive FOBT, and after colonoscopy had been performed. As the FOBT had already been performed and classed all the samples as 'positive', then by carrying out the analysis following this, any ability to separate the samples in the correct groups demonstrates an improvement over the FOBT screening method.

Furthermore, both univariate and multivariate data analysis of the SIFT-MS data enabled specific compounds to be tentatively identified that were significantly different between the low risk and high risk groups. One ion was m/z 35 using H_3O^+ which is likely to be hydrogen sulfide. As discussed earlier, it has been well established that hydrogen sulfide plays a major role in the colon and may be implicated in colorectal cancer. This may indicate a change in gut flora, and it is possible that sulfate reducing bacteria are processing increased available hydrogen, or in fact that the sulfate reducing bacteria are present in a greater amount in the bowel of

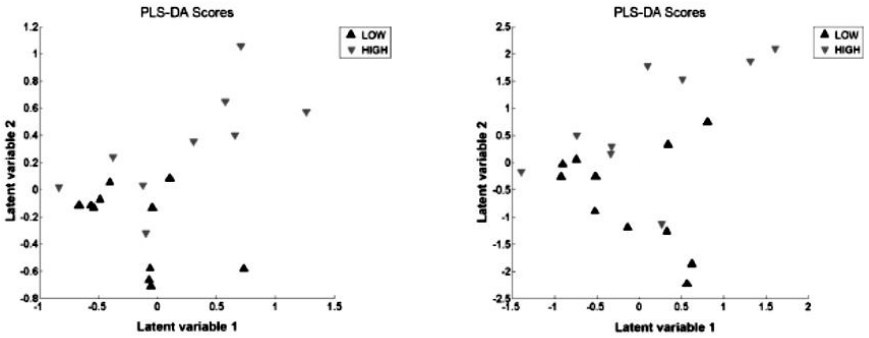


Fig 5. PLS-DA scores which show distinction between low and high risk extracted from the optimal models following feature selection via STT (A) and WTT (B).

doi:10.1371/journal.pone.0130301.g005

the high risk group. Shen et al. [43] investigated adherent bacteria in the gut mucosa of adenoma and non-adenoma subjects. They found a significant higher abundance of Proteobacteria and significantly lower abundance of Bacteriodetes in the adenoma patients than the non-adenoma. Sulfate reducing bacteria are found in the 5 distinct genera of the Delta subdivision of Proteobacteria phylum. Three of these genera consume partly reduced fermentation products (e.g. lactate) and reduce sulfate to sulfide, whilst the other two are hydrogen consuming [37, 44]. Red meat is also thought to increase the levels of sulfate producing bacteria which are able to use the sulfur residues from meat [45]. Higher levels of fat, red meat, and obesity have all been linked to an increased risk of colorectal cancer so could all be very relevant.

For m/z 35, the counts per second for the high risk group was 965 which had a much broader range than the low risk group, along with the higher count rate, suggesting that the m/

Table 2. m/z values suggested by the PLS-DA loadings when combining the precursor ions together for analysis using PLS-DA following feature selection with STT and WTT.

STT		WTT	
M/z Value and precursor ion	**Possible Compound(s)**	**M/z Value and precursor ion**	**Possible Compound(s)**
m/z 18—H_3O^+, m/z 54—H_3O^+	Ammonia	m/z 18—H_3O^+	Ammonia
m/z 35—H_3O^+	Hydrogen Sulphide	m/z 20—H_3O^+	unknown
m/z 53—H_3O^+ m/z 62—H_3O^+	Unknown	m/z 35—H_3O^+	Hydrogen Sulphide
m/z 68—H_3O^+	Pyrrole	m/z 53—H_3O^+	Unknown
m/z 80—H_3O^+	Pyridine	m/z 81—H_3O^+	Acetaldehyde
m/z 90—H_3O^+ m/z 108—H_3O^+	Unknown	m/z 90—H_3O^+	Unknown
m/z 115—H_3O^+	Heptanal	m/z 111—H_3O^+	Octanal; propanoic acid
m/z 44—NO^+	Isopropylamine, Dimethylamine, Methylethylamine	m/z 115—H_3O^+	Heptanal
m/z 103—NO^+	Unknown	m/z 44—NO^+	Isopropylamine, Dimethylamine, Methylethylamine
m/z 52—O_2^+ m/z 60—O_2^+ m/z 70—O_2^+ m/z 109—O_2^+ m/z 123—O_2^+	Unknown	m/z 104—NO^+	Propanoic acid, methyl acetate, ethyl formate
		m/z 50—O_2^+ 52—O_2^+ m/z 60—O_2^+ m/z 61—O_2^+ m/z 70—O_2^+ m/z 109—O_2^+	unknown

doi:10.1371/journal.pone.0130301.t002

z 35 was far more prevalent in the high risk samples. This m/z value was also identified when using multivariate analysis as a significant m/z value which further strengthens the likelihood that it may be clinically relevant. In addition, when the O_2^+ data were assessed using univariate analysis, two ions, m/z 62 and m/z 94 were identified which could well prove to support the theory that increased production of sulphides is more prevalent in the high risk group as they were identified as the compounds dimethyl sulphide and dimethyl disulphide respectively.

The other two compounds that were responsible for separating the groups using multivariate statistics were ammonia and acetaldehyde. M/z 18 and 54 were significant ions; m/z 18 being NH_4^+, and m/z 54 being $NH_4^+.2H_2O$. These ions were higher in the high risk group than the low risk group. As previously mentioned, ammonia is the product when undigested protein that reaches the colon is fermented by the microflora. High levels of ammonia have been reported to have a range of toxic effects which include enhancing cell proliferation and favouring the growth of malignant cells [46]. When all three precursors were analysed combined, none of the acetaldehyde m/z values were found to be significant with PLS-DA and STT but m/z 81 (which is protonated acetaldehyde with 2 water molecules) was significant using PLS-DA and WTT. The potential importance of this compound in colorectal cancer is that it is a breakdown product of alcohol (ethanol). Alcohol is shown to be a risk factor for colorectal cancer [47], and if increased amounts of alcohol are being produced, or if bacteria in the gut are producing increased volumes of alcohol—subsequently leading to increased acetaldehyde—then it is likely this will cause damage. Acetaldehyde is broken down by alcohol dehydrogenase and is usually quickly metabolised by an aldehyde dehydrogenase to acetate. However in the gut, it can locally accumulate when the local microbiota oxidise ethanol and there, if folate is present, the acetaldehyde degrades it and can cause folate deficiency [48]. These results would suggest that either alcohol consumption should be reduced, or again that the balance of microbiota needs to be assessed to rebalance the breakdown of products that may lead to adverse effects.

Other compounds tentatively identified include pyridine and pyrrole. When looking at pyridine alone, no reported link is found with colorectal cancer, however the heterocyclic amine (HCA) 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) has been widely identified as implicated in the aetiology of human colorectal cancer. Heterocyclic amines are formed during the cooking of protein rich foods such as meat, especially by high temperature methods (above 150°C) such as grilling, barbequing and pan frying. PhIP is seen to be of greater significance in regards to colorectal cancer because it is the predominant HCA found in cooked meats [49,50]. Heating meat creates various imidazoquinoline, imidazoquinoxaline and imidazopyridine compounds which are potent and highly mutagenic towards some strains of Salmonella typhimurium [51]. Conversely some pyrrole pigments have been identified that have carcinogen adsorption properties, which include hemin and chlorophyllin. In some studies antimutagenesis has been found with these and other pyrrole pigments both in vitro and in vivo [52].

One note of caution in dealing with the analysis of VOCs from fecal samples is that because patients produce fecal samples in their own time, and not always at clinic (as would be the case with urine or blood samples, for example), there is some uncertainty as to how samples are handled and stored prior to their return to the clinic/lab and then appropriate storage. Patients may be given advice on the appropriate way of handling and storing samples, but researchers have no way of knowing whether this has been done. This may cause variability in the quality of samples and thus reduce the effectiveness of this otherwise promising technique.

Conclusions

Overall, what all the data suggest is that a difference can be predicted between the high risk group and the low risk group just by employing the multivariate model between the known

high risk and low risk samples. This model could then be employed on unknown samples to predict if they would likely fall into the high risk or low risk group, based on their overall 'metabolic' SIFT-MS profile. This could then be useful in predicting whether someone is going to have or get colorectal cancer, and this method is likely to produce far fewer false positives than FOBT, thus reducing the number of unnecessary colonoscopies.

What could also then be added to strengthen the overall clinical picture is that the sample could then be analysed to specifically look for the compounds discussed and if these presented above a certain concentration, they would all suggest further investigation may be more valid for colorectal cancer. In particular, hydrogen sulphide may be a significant marker to aid in the screening or diagnosis of colorectal cancer.

Supporting Information

S1 Table. A table of the SIFT-MS data used to generate the findings reported here is available in Microsoft Excel format.
(XLSX)

Acknowledgments

The authors are grateful to Mr Nigel Hall and the department of colorectal surgery at Addenbrooke's Hospital for their support.

Author Contributions

Conceived and designed the experiments: CT JOH. Performed the experiments: CAB MC CL. Analyzed the data: CAB MC CL CT. Contributed reagents/materials/analysis tools: JOH. Wrote the paper: CAB MC JOH CT.

References

1. Center MM; Jemal A; Smith RA; Ward E. (2009a)—Worldwide Variations in Colorectal Cancer, CA: A Cancer Journal for Clinicians, 59, 366–378. doi: 10.3322/caac.20038 PMID: 19897840
2. Wu AH; Peganini-Hill A; Ross RK; Henderson BE. (1987)—Alcohol, physical activity and other risk factors for colorectal cancer: A prospective study, British Journal of Cancer, 55, 687–694 PMID: 3620314
3. Center MM; Jemal A; Ward E. (2009b)–International Trends in Colorectal Cancer Incidence Rates–Cancer Epidemiology, Biomarkers & Prevention, 18, 1688–1694.
4. Lieberman DA. (2009)–Screening for Colorectal Cancer, New England Journal of Medicine, 361, 1179–1187 doi: 10.1056/NEJMcp0902176 PMID: 19759380
5. Rhodes JM. (2000)—Colorectal cancer screening in the UK: Joint Position Statement by the British Society of Gastroenterology, The Royal College of Physicians, and The Association of Coloproctology of Great Britain and Ireland, Gut, 46(6), 746–748. PMID: 10807879
6. UK Colorectal Cancer Screening Pilot Group—BMJ, doi: 10.1136/bmj.38153.491887.7C (published 5 July 2004)
7. Weller D; Coleman D; Robertson R; Butler P; Melia J; Campbell C et al. (2007)–The UK Colorectal Cancer Screening Pilot: Results of the Second Round of Screening in England, British Journal of Cancer, 97, 1601–1605 PMID: 18026197
8. NHS. (2014)–NHS Bowel Cancer Screening Programme–[Online]–Accessed May 2014 –Available from: http://www.cancerscreening.nhs.uk/bowel/
9. Vernon SW. (1997)–Participation in Colorectal Cancer Screening: A Review, Journal of the National Cancer Institute, 89(19), 1406–1422 PMID: 9326910
10. Winawer S; Fletcher R; Rex D; Bond J; Burt R; Ferrucci J et al.; Gastrointestinal Consortium Panel. (2003)–Colorectal Cancer Screening and Surveillance: Clinical Guidelines and Rationale—Update Based on New Evidence, Gastroenterology, 124, 544–560 PMID: 12557158
11. Levin B; Lieberman DA; McFarland B; Andrews KS; Brooks D; Bond J et al.; American Cancer Society Colorectal Cancer Advisory Group; US Multi-Society Task Force; American College of Radiology

Colon Cancer Committee. (2008)—Screening and Surveillance for the Early Detection of Colorectal Cancer and Adenomatous Polyps, 2008: A Joint Guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology–CA: A Cancer Journal for Clinicians, 58, 130–160 doi: 10.3322/CA.2007.0018 PMID: 18322143

12. Strul H & Arber N. (2002)–Faecal Occult Blood Test for Colorectal Cancer Screening, Annals of Oncology, 13, 51–56 PMID: 11863111
13. Department of Health (DOH) (2012)–Bowel Cancer Screening–The Facts, Cancer Research UK & the NHS, DH Publications
14. Towler B; Irwig L; Glasziou P; Kewenter J; Weller D; Silagy C. (1998)–A Systematic Review of the effets of Screening for Colorectal Cancer using the Faecal Occult Blood Test, British Medical Journal, 317, 559–565 PMID: 9721111
15. Mandel JS; Church TR; Bond JH; Ederer F; Geisser MS; Mongin SJ et al; (2000)–The Effect of Fecal Occut Blood Screening on the Incidence of Colorectal Cancer, New England Journal of Medicine, 343, 1603–1607 PMID: 11096167
16. Yeasmin F; Rahman MA; Ali A; Sultana T; Rahman MdQ; Ahmed ANN et al. (2011)—Role of Immunological Method of Fecal Occult Blood Test for Screening Colorectal Diseases, Bangabandhu Sheikh Mujib Medical University Journal, 4(2), 76–80
17. Lowenfels AB. (2006)–Fecal Occult Blood Testing as a Screening Procedure for Colorectal Cancer, Annals of Oncology, 13, 40–43
18. Imperiale TF; Ransohoff DF; Itzkowitz SH; Turnbull BA; Ross ME; (2004) Fecal DNA versus Fecal Occult Blood for Colorectal-Cancer Screening in an Average-Risk Population for the Colorectal Cancer Study Group* N Engl J Med 351, 2704–14 PMID: 15616205
19. Marchesi JR; Dutilh BE; Hall N; Peters WHM; Rian Roelofs R; Boleij A et al.; (2011)—Towards the Human Colorectal Cancer Microbiome, PLOS One, 6(5), 1–8.
20. Zimmermann D; Hartmann M; Moyer MP; Nolte J; Baumbach JI. (2007)—Determination of volatile products of human colon cell line metabolism by GC/MS analysis, Metabolomics, 3 (1), 13–17
21. Denkert C; Budczies J; Weichert W; Wohlgemuth G; Scholz M; Kind T et al.; (2008)—Metabolite profiling of human colon carcinoma—deregulation of TCA cycle and amino acid turnover, Molecular Cancer, 7(72), 1–15 doi: 10.1186/1476-4598-7-72 PMID: 18799019
22. Willis CM; Church SM; Guest CM; Cook WA; McCarthy N; Bransbury AJ et al.; (2004)—Olfactory detection of human bladder cancer by dogs: proof of principle study, British Medical Journal, 329(7468), 712. PMID: 15388612
23. Sonoda H; Kohnoe S; Yamazato T; Satoh Y; Morizono G; Shikata K et al.; (2011)—Colorectal cancer screening with odour material bycanine scent detection, Gut, 60, 814–819 doi: 10.1136/gut.2010.218305 PMID: 21282130
24. Altomare DF; Di Lena M; Porcelli F; Trizio L; Travaglio E; Tutino M et al; (2013)–Exhaled Volatile Organic Compounds Identify Patients with Colorectal Cancer, British Journal of Surgery, 100, 144–150. doi: 10.1002/bjs.8942 PMID: 23212621
25. Probert CSJ; Ahmed I; Khalid T; Johnson E; Smith S; Ratcliffe N. (2009)–Volatile Organic Compounds as Diagnostic Biomarkers in Gastrointestinal and Liver Diseases, Journal of Gastrointestinal and Liver Disease, 18(3) 3, 337–343. PMID: 19795029
26. Cauchi M, Fowler DP, Walton C, Turner C, Jia W, Whitehead RN et al; (2014), Application of gas chromatography mass spectrometry (GC-MS) in conjunction with multivariate classification for the diagnosis of gastrointestinal diseases, Metabolomics, doi: 10.1007/s11306-014-0650-1
27. Smith D; ŠpanÆl P. (2005)–Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) for online trace gas analysis, Mass Spectrometry Reviews, 24, 661–700 PMID: 15495143
28. Nachr N. (2008)–The Mann-Whitney U: A Test for assessing whether two independent samples come from the same distribution, Tutorials in Quantitative Methods for Psychology, 4(1), 13–20.
29. Wold S., Esbensen K., and Geladi P., Principal component analysis. Chemometrics and Intelligent Laboratory Systems, 1987. 2(1–3): p. 37–52
30. Student THE PROBABLE ERROR OF A MEAN. Biometrika, 1908. 6(1): p. 1–25
31. Wilcoxon F., Individual comparisons by ranking methods. Biometrics Bulletin, 1945. 1(6): p. 80–83
32. Barker M. and Rayens W., Partial least squares for discrimination. Journal of Chemometrics, 2003. 17 (3): p. 166–173
33. Brereton R.G., *Chemometrics for Pattern Recognition* 2009, Chichester, UK: John Wiley & Sons
34. Campbell M.J. and Machin D., *Medical Statistics: A Common Sense Approach*. 3rd Edition ed. 1999, Chichester, UK.: John Wiley & Sons Ltd

35. Medani M; Collins D; Docherty NG; Baird AW; O'Connell PR; Winter DC (2011)—Emerging Role of Hydrogen Sulfide in Colonic Physiology and Pathophysiology, Inflammatory Bowel Disease, 17,1620–1625 doi: 10.1002/ibd.21528 PMID: 21674719

36. Picton R; Eggo MC; Merrill GA; Langman MJS; Singh S. (2002)—Mucosal protection against sulphide: importance of the enzyme rhodanese, Gut, 50, 201–205 PMID: 11788560

37. Gibson GR; McFarlane GT, Cummings JH. (1993)—Sulphate reducing bacteria and hydrogen metabolism in the human large intestine, Gut, 34, 437–439 PMID: 8491386

38. Li L; Rose P; Moore PK. (2011)—Hydrogen Sulfide and Cell Signaling, Annual Review of Pharmacology and Toxicology, 51, 169–187 doi: 10.1146/annurev-pharmtox-010510-100505 PMID: 21210746

39. Attene-Ramos MS; Wagner ED; Plewa MJ; Gaskins HR. (2006)—Evidence That Hydrogen Sulfide Is a Genotoxic Agent, Molecular Cancer Research, 4(1), 9–14 PMID: 16446402

40. Cai WJ; Wang MJ; Ju LH; Wang C; Zhu YC. (2010)—Hydrogen sulfide induces human colon cancer cell proliferation: role of Akt, ERK and p21, Cell Biology International, 34, 565–572 doi: 10.1042/CBI20090368 PMID: 20184555

41. Jia W; Whitehead R; Griffiths L; Bai H; Waring R; Ramsden D et al.; (2012)–Diversity and distribution of sulphate-reducing bacteria in human faeces from healthy subjects and patients with inflammatory bowel disease, FEMS Immunology and Medical Microbiology, 65, 55–68. doi: 10.1111/j.1574-695X.2012.00935.x PMID: 22309113

42. Ramasamy S; Singh S; Taniere P; Langman MJS; Eggo MC. (2006)—Sulfide-detoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation—American Journal of Physiology, Gastrointestinal and Liver Physiology, 291, G288–G296. PMID: 16500920

43. Shen X, Rawls JF; Randall T; Burcall L; Mpande CN; Jenkins N. et al. (2010)—Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas—*Gut microbes* 1(3) 138–147.

44. Dethlefsen L., Eckburg PB; Bik EM; Relman DA. (2006)—Assembly of the human intestinal microbiota —Trends in Ecology & Evolution 21(9) 517–523.

45. Ou J; Carbonero F; Zoetendal EG; DeLany JP; Wang M; Newton K et al.; (2013)—Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans.–American Journal of Clinical Nutrition, 98(1), 111–120 doi: 10.3945/ajcn.112.056689 PMID: 23719549

46. Birkett A; Muir J; Phillips J; Jones G; O'Dea K. (1996)—Resistant starch lowers fecal concentrations of ammonia and phenols in humans, The American Journal of Clinical Nutrition, 63, 766–772 PMID: 8615362

47. Bardou M; Montembault S; Giraud V; Borotto E; Houdayer C; Capron F et al.; (2002), Excessive alcohol consumption favours high risk polyp or colorectal cancer occurrence among patients with adenomas: a case control study, *Gut*, 50:38–42 doi: 10.1136/gut.50.1.38

48. Homann N; Tillonen J; Salaspuro M. (2000)–Microbially Produced Acetaldehyde from Ethanol may Increase the risk of Colon Cancer via Folate Deficiency, International Journal of Cancer, 86, 169–173 PMID: 10738242

49. Coles B; Nowell SA; MacLeod SL; Sweeney C; Lang NO; Kadlubar FF. (2001)–The role of human gluththione S-transferases (hGSTs) in the detoxification of the food-derived carcinogen metabolite N-acetoxy-PhIP, and the effect of polymorphism in hGSTA1 on colorectal cancer risk–Mutation Research, 482, 3–10. PMID: 11535243

50. Miller PE; Lazarus P; Lesko SM; Cross AJ; Sinha R; Laio J et al. (2013)–Meat-Related Compounds and Colorectal Cancer Risk by Anatomical Subsite–Nutrition Cancer, 65(2), 202–226. doi: 10.1080/01635581.2013.756534 PMID: 23441608

51. Nakagama H; Nakanishi M; Ochiai M. (2005)–Modelling human colon cancer in rodents using a food-borne carcinogen, PhIP–Cancer Science, 96(10), 627–636. PMID: 16232193

52. Ferguson LR; Philpott M; Karunasinghe N. (2004)–Dietary Cancer and prevention using antimutagens, Toxicology, 198, 147–159. PMID: 15138038